TAMPERE UNIVERSITY OF TECHNOLOGY

AMEDEO PISAPIA

SMOOTHING OF PHOTOVOLTAIC POWER PRODUCTION BY PLANT DISPERSION

Master's Thesis

Examiner: Professor Seppo Valkealahti

The examiner and topic of the thesis were approved by the Council of the Faculty of Computing and Electrical Engineering on 6 April 2016

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Electrical Engineering
AMEDEO PISAPIA: Smoothing of Photovoltaic power production by plant dispersion
Master of Science Thesis, 80 pages
August 2017
Major: Renewable Energy
Examiner: Seppo Valkealahti
Keywords: Smoothing effect, PV plant dispersion, distributed PV power plant

Power production by a (Photovoltaic) PV power plant is strongly dependent on the irradiance level that is incident on the panels thereof. By neglecting the sun's variability that is precisely predictable, the main factor that influences the output power is the movement of the clouds over the plant. Therefore, such moving clouds lead to the presence of some fluctuations on the output power, that may influence the power quality and cause some problems of matching between the PV source and the load. In order to compensate such fluctuations, some Energy Storage Systems (ESS) are commonly used, but they need to be sized correctly. Many studies demonstrate that the solar variability impact on produced power is different if we consider a set of geographically dispersed systems, instead of a single system. It has been noted a strong reduction in variability when the aggregation of several PV systems is taken into account, rather than just one single PV system. That is called "smoothing effect".

In order to evaluate the output power of a single PV system, it was shown in literature that a PV system can be seen as a low-pass filter, due to its physical extension, where the input is the measured irradiance at a single point, and the output is a smoothed version of the single point irradiance, such filter output is proportional to the output power.

Other studies carried out how to evaluate the smoothing effect due to the geographical dispersion of a distributed power plant. Using these methods, it is possible quantifying the reduction of variability associated with the dispersion of the plant over a region.

After having developed and validated a method to estimate the variability of a distributed power plant, the main objective of this master thesis has been to analyse the power output variability of a PV power plant by using the irradiance data measurements done in Tampere (Finland), considering different scenarios by changing the capacity and the number of subsystems that composes the power plant. This is done in order to compare the different layouts and understand how the smoothing effect depends on the plant size and on the geographical dispersion.

Preface

This Master of Science Thesis has been developed, during the Erasmus experience, at the Deparment of Electric Energy Engineering of the Tampere University of Technology (TUT). The supervisor of my thesis work has been Prof. Seppo Valkealahti.

First of all, I want to thank the Professor Seppo Valkealahti for his guidance over this experience, he has been very helpful and patient over these last months.

I am grateful to my parents for their support during these years, they have always believed in me during all my life.

I would like to say thank you to all of the people that I have met during my Erasmus at TUT.

A particular mention is deserved by my friend and roommate Nico, the funny Francesco, and the crazy Pino'.

Lastly, I would like to mention the wonderful women that I have met during this experience, the delicate and shy Michelle, the caring and confidant Rosanna, and the pretty and cheerful Sara, I feel so lucky to have met you!!

Kiitos, Suomi!

Amedeo Pisapia

Table of contents

List of Figures

List of Tables

Symbols and Abbreviations

Notation

A	Correlation scaling factor
C	Capacity
D	Dispersion Factor
$d_{m,n}$	distance between locations m and n
$E_{\Delta t,d}$	Daily maximum power fluctuation error
$F(t)$	Fluctuant Irradiance
f_c	Cut-off frequency
fl	Fluctuation Factor
fpi	fluctuation power index
$G(t)$	Irradiance Data
$H(s)$	Transfer Function
L	Length of the power plant
$P(t)$	Output Power
$\Delta P(t)$	Output Power Variability
Δt	Time interval
$\rho(d_{m,n}, \bar{t})$	Correlation factor
$\sigma_{\Delta t}^{Fleet}$	Standard deviation of the output power changes
$\sigma_{\Delta t}^{Distributed}$	Standard deviation of the output power changes of a distributed power plant
$\sigma_{\Delta t}^{system}$	Standard deviation of the output power changes of a single location
S	Plant Area
$\bar{t}$	Timescale of fluctuations
v	Wind velocity
Var	Variance
VR	Variability Reduction
$\omega_{\bar{t}}(t)$	Mode of the fluctuation

Abbreviations

cs	Cloud speed
DFT	Discrete Fourier Transform
ESS	Energy Storage System
FFT	Fast Fourier Transform
N	Number of subsystems
PDF	Probability Density Function
PV	Photovoltaic
RMSE	Root Mean Square Error
ROV	Relative Output Variability
STC	Standard Test Condition
WVM	Wavelet Variability Model

1 Introduction

A “PV power plant” is a facility for the generation of electric power from the solar energy. Such power plant can be classified in two categories depending on how it is built:

- Centralized system, that is just a single PV system.
- Distributed fleet composed by more PV (sub)systems, located in different locations, connected to the same electrical grid. N denotes the number of subsystems which composes the fleet.

A “PV system” is an aggregation of more PV arrays located approximately in the same location.

The nominal output power of a PV system is the power that the source can supply to the load under Standard Test Condition (STC), so with an incident irradiance equal to 1000 W/m^2 and a temperature of 25°C. This quantity is specified with the term “capacity” and is commonly denoted with C.

The variability of the output power (commonly named “step change” too) $\Delta P(t)$, at an instant *t* for a given sampling period Δt, is the difference between two adjacent power output samples normalized to the capacity C of the plant under consideration and then multiplied by 100 to have the results in percentage, as reported in equation (1).

$$\Delta P(t) = \frac{[P(t+\Delta t) - P(t)]}{C} \times 100 \tag{1}$$

For example, a ΔP equal to 20%, on a power plant with a capacity C of 1 MW, means that a fluctuation of 0.2 MW occurred over a single Δt.

The study of the solar resource variability has been becoming an important topic in the last decades because of the increasing penetration of solar photovoltaic panels on the electrical grid. This kind of variability may lead to voltage changes on the distribution system and have a negative impact on power quality.

Therefore, the main objective of the research in this field is to quantify and reduce these fluctuations, this is done mainly in order to size correctly the Energy Storage Systems.

PV power output depends on the irradiance incident on the panels, changes in solar irradiance depend on the sun and the movement of the clouds; while the sun's variability is precisely predictable, the variability caused by the moving clouds needs to be studied.

Some studies demonstrate that the solar variability impact on produced power is different if we consider a set of geographically dispersed systems, instead of a single system. It has been demonstrated a strong reduction in variability when the aggregation of several PV systems is taken into account, rather than just one single PV system. That is called "smoothing effect".

In order to evaluate the output power of a single PV system, it was shown in literature that a PV system can be seen as a low-pass filter, due to its physical extension, where the input is the measured irradiance at a single point, and the output is a smoothed version of the single point irradiance, such filter output is proportional to the output power.

Other studies carried out how to evaluate the smoothing effect due to the geographical dispersion of a distributed power plant. Using these methods, it is possible quantifying the reduction of variability associated with the dispersion of the plant over a region.

After having developed and validated a method to estimate the variability of a distributed power plant, the main objective of this master thesis is to analyse the power output variability of a PV power plant by using the irradiance data measurements done in Tampere (Finland), considering different scenarios by changing the capacity and the number of subsystems that composes the power plant. This is done in order to compare the different layouts and understand how the smoothing effect depends on the plant size and on the geographical dispersion.

The thesis is structured as follows:

In the chapter 2, the most important studies on the output power variability and the "smoothing effect" are shown and discussed, firstly by observing the power and

irradiance data, and then by showing the methods proposed in literature which estimate and quantify such variability.

In the chapter 3, it is shown how the calculation of the power output variability is performed, with a special focus on the development and validation of a method to simulate the output power of a distributed power plant, and therefore estimate its variability.

In the chapter 4, a comparison among different scenarios is proposed. Many metrics are employed in order to figure out which solution provides a more stable output power and try to understand how the plant size and the geographical dispersion influence the output power variability.

2 State of the art of PV power variability studies

In this section, it will be introduced one of the main problems of PV applications, that is source variability, and presented an overview over some results and some existing methods for estimating the power variability in PV systems.

Unlike conventional electrical power generation (e.g., fossil or nuclear), solar energy is intermittent. Since the output power of any PV system depends on the irradiance incident on the panels with a quite proportional law, the output of a solar power plant is driven by the cycle of days and seasons and by the weather conditions.
Therefore, the intermittency, or better termed, variability, of the solar resource has two components:

- The former is precisely predictable and traceable due to the apparent motion of the sun in the sky and the earth's distance from the sun.
- The latter, conversely, is much less predictable and it is due to the motion of clouds.

These changes in solar irradiance occur in a wide range of timescales, from few milliseconds to several hours. As we can see in Table 1, each of these changes will cause a different kind of impact on the power system [1].

Table 1: Potential power system impacts of solar irradiance variability [1].

Timescale of changes in solar irradiance	Potential power system impact
Seconds	Power quality (e.g. voltage flicker)
Minutes	Regulation reserves
Minutes to hours	Load following power plant

In particular, the changes on seconds timescale may influence the power quality, specifically having some problems of compatibility between the output of the PV source and the load, this is caused by sudden increase or decrease in voltage.
Over longer timescales, the variability concerns the auxiliary systems of the electrical grid such as the regulation reserves, that are reserve units of energy capable of compensate voltage imbalances caused by the random nature of loads, and the load

following power plants that adjust their power output to follow changes required by the load throughout a day.

The typical timescale of the sun's variability ranges between minutes and hours and is caused by the sun's trajectory, which is precisely predictable, so it is not interesting to study it. It is more interesting a focus on the shorter variability that is induced by the movement of clouds.

A common way to compensate the fluctuations that we have, it is using some Energy Storage Systems (ESS). These systems are quite expensive and so their requirements need to be sized correctly. Underestimating these requirements may result to an unstable power system and overestimating reserve requirements leads to a useless extra cost, so it is very important to be able to quantify the phenomenon of power output variability.

To quantify correctly the amount of output fluctuations of a photovoltaic power plant, an important effect needs to be considered. Many studies showed that the aggregated output of many photovoltaic power systems geographically distributed over a wide area fluctuates less than the output of the individual systems, this effect is known like "smoothing effect". Many methods have been presented in the scientific literature to estimate the power output variability considering this effect.

In the paragraph 2.1, some studies, with the aim to points out some conclusions on the smoothing effect by observing the power output data, are shown. The same is done in the paragraph 2.2, where a similar data analysis is deployed by observing the irradiance data instead of the power ones. In the paragraph 2.3, a description of some common methods used to estimate the power output variability is reported. At the end, in the paragraph 2.4, it is described how the contents presented along this chapter are used.

2.1 Observed variability and smoothing in monitored PV power production data

Three studies, carried out in two years between 2010 and 2011 from the American company SunEdison, show empirically how the variability of the PV system ensembles output power depends on the sizes of individual systems and on the geographical dispersion of systems. In [2] three PV power plants located in California and New Jersey were studied. In total 67 PV systems were considered in the study, the power production data were measured with a 1-min resolution during seven days in May chosen for the analysis. The PV systems were grouped in ensembles with different total capacities and some different compositions of systems for each total capacity, in order to quantify the output variability of a single PV system or an ensemble of PV systems they used three different metrics:

- The standard deviation of the output power changes ΔP between consecutive minutes
- The maximum drop (or "most negative ΔP") between any two consecutive minutes
- The third is the percentage of ΔP values that are higher than 10% of the power plant capacity within the analysed interval, this quantity gives an idea of how many fluctuations can be considered significant.

Each ensemble consists of many small or a few large PV systems giving the same total capacity within roughly the same area. More systems imply a higher fragmentation over the service territory, but no explicit relation to the degree of dispersion was evaluated. The main finding is that power production restricted to one or a few sites is more variable than production dispersed on a large number of smaller systems. Table 2 puts in evidence the difference in maximum (negative) step change between the least and most dispersed PV ensembles for each ensemble capacity. The differences are dependent not only on the number of systems aggregated, but also on the composition of the ensemble (whether a few systems are dominating in terms of capacity or if there is an even distribution), the exact area for dispersion, and the irradiation variability on the actual day. Nevertheless, Table 2 shows the amount of the smoothing that can be achievable through distribution of PV systems. By distributing

the same capacity over a larger area, the maximum 1-min step change decreases from around 40-50% of capacity to around 5% of capacity. Note, however, that since only one day was studied for each ensemble, the observed maximum step change is probably lower than what would appear over a longer observation period.

Table 2: Maximum PV step changes during one observed day in [2]. Note that the degree of dispersion depends both on the number of systems and whether one or a few systems are dominating, which is why the least dispersed ensembles can contain quite a few systems [3].

	Least dispersed ensemble		**Most dispersed ensemble**	
Ensemble size (kW_p)	**Maximum 1-min step change (%)**	**Number of systems**	**Maximum 1-min step change (%)**	**Number of systems**
440	37	1	6	11
1000	17	15	8	17
1200	50	1	17	8
1200	43	2	5	15
2300	36	2	18	11
2500	21	20	8	22

This is apparent in [4], where the output of distributed PV systems in one service area in New Jersey composed by 31 systems was analysed over the longer period of 11 months with the same approach and considering three different time scales: 1-minute, 10-minute and 60-minute. The systems were grouped in seven differently composed ensembles, four groups with capacity equal to 440 kW_p and three groups with capacity equal to 1000 kW_p. The main result is naturally that the variability metrics increase systematically with increasing aggregation of individual systems, while the dependence on the time resolution is less coherent. Fig. 1 and Fig. 2 show the maximum step change for all seven ensembles. It is clear that, when the capacity is spread over a larger number of systems, the maximum step change is smaller (S11 in Fig. 1 and L11 in Fig. 2).

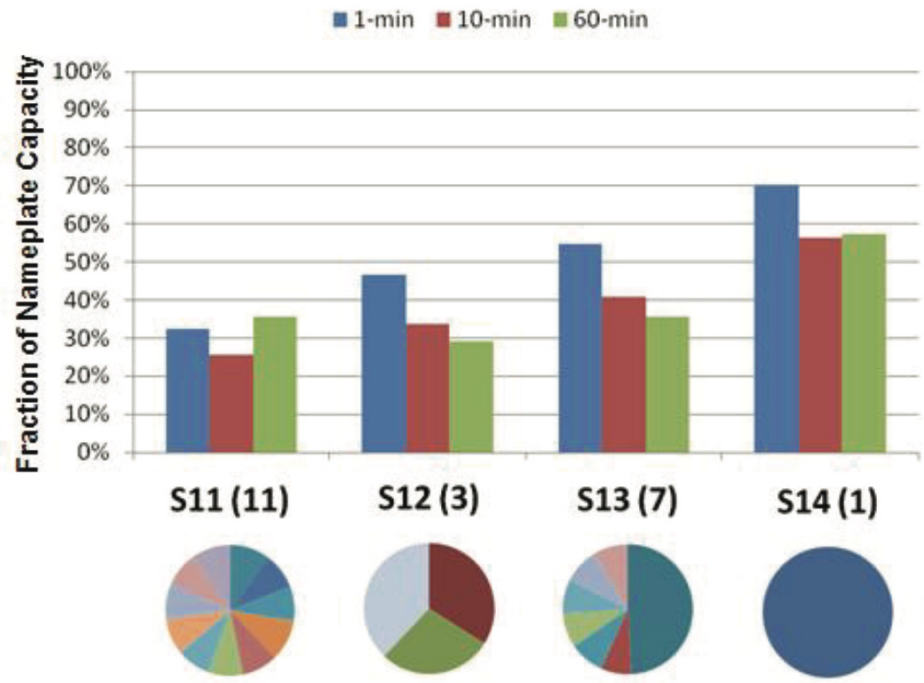

Fig. 1: Maximum step change for ensembles sizing 400 kW_p for all time scales, plotted as a fraction of the aggregate nameplate capacity. The pie charts display the relative size of the systems comprising the ensemble [4].

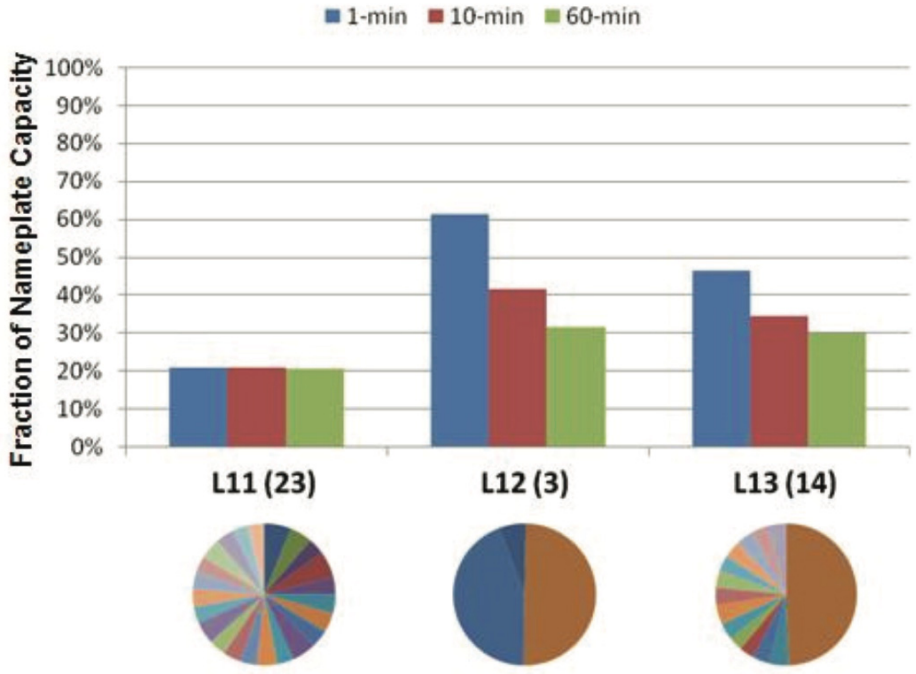

Fig. 2: Maximum step change for ensembles sizing 1000 kW_p for all time scales, plotted as a fraction of the aggregate nameplate capacity. The pie charts display the relative size of the systems comprising the ensemble [4].

These maximum step changes are higher than in the previous study, this can be explained by considering that the observation period is longer (11 months vs. 7 days). Since an observed maximum is dependent on the observation period, a better way to

compare the results is using a percentile of the distribution of the step changes, the 95^{th} and 99.7^{th} percentiles are commonly used. This is done in order to leave the more extreme values out of the analysis, such values are omitted because over a longer observation period the maximum step change can be very rare, so a metric as the 99.7^{th} percentiles is fairer. By analysing the step changes, for the most dispersed ensemble (L11 in Fig. 2) in this study the 95^{th} percentile is 5% of power plant capacity and the 99.7^{th} percentile is 11%, let us remember that the maximum step change was about 20%.

In [5], the same authors investigated the output fluctuations from a centralized PV plant that was Europe's largest at the time of writing (2011). This plant, located in Rovigo, Italy, has a capacity of 70 MW_p in total and consists of 60 individual arrays, covering a total area of 850.000 m^2. In addition to the previous two studies, the authors systematically analysed the impact of dispersion on variability (over relatively short distances, at maximum 1200 m between systems), as compared to the impact from capacity increases within the same limited area. This was done by on the one hand grouping the arrays into differently sized ensembles (8, 17, 35 and 70 MW_p) within the same area, and on the other hand considering ensembles of the same size (6 MW_p) dispersed on circles with different radii. The result was, naturally, that increased ensemble size dispersed within the same area does not decrease variability (by considering time scales from 1-minute to 60-minute), whereas wider dispersion of the same capacity does as shown, respectively, in Fig. 3 and Fig. 4.

Furthermore, dispersion of capacity along circles with radii from 100 to 600 m made the maximum observed step change over 9 months decrease from almost 70% to around 45% of capacity at the 1-min resolution. Note the difference to the previous studies, where the wider dispersion (maximum distances between systems in the order of 100 times longer) gave considerably lower maximum step changes.

This puts in evidence the importance of the distance between the PV systems, longer distances lead to a more stable output power.

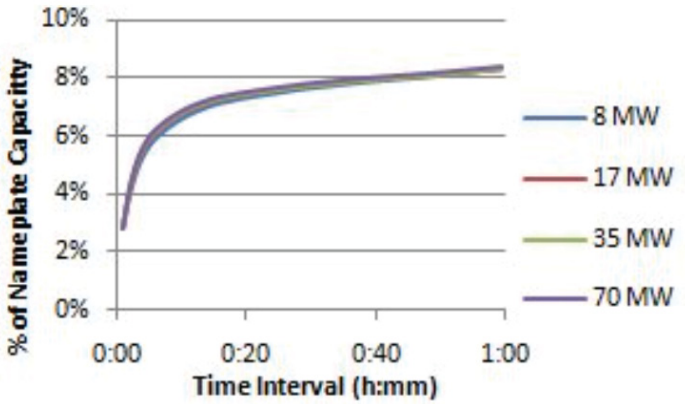

Fig. 3: Standard Deviation of four array ensembles within the same area by varying capacity, the result does not change for all time scales [5].

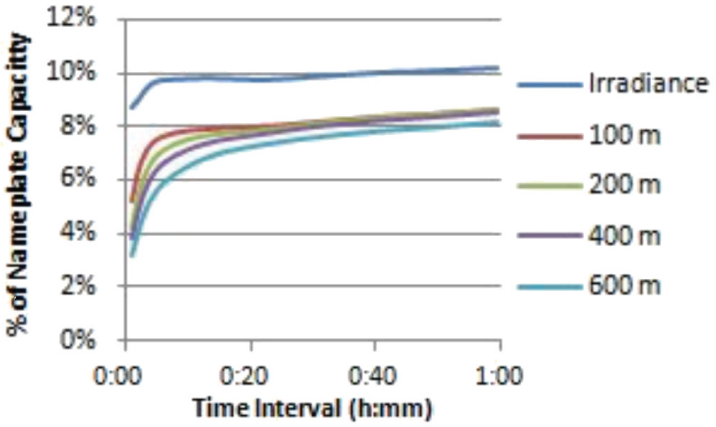

Fig. 4: Standard Deviation of four array ensembles (plus zero-radius measurement) with same capacity by varying radii, the result changes for all time scales [5].

In [6], it is illustrated the impacts of transient clouds on the energy production at a large PV plant covering 728.000 m^2 (180 acres) of land with a rated capacity of 25 MW_p, located in Florida. The analysis is based on electricity production data monitored at the 17 subsystems of the plant. The results presented considers only one day in December, but with a remarkable time resolution of 10 seconds, which is higher than the ones in the studies above. The impact of dispersion over the area of the plant can be clearly seen. The maximum 10-second step change observed over this day with passing clouds is around 60 % of capacity for individual subsystems (sized 0.8-1.6 MW_p) and around 5% for the whole 25 MW_p system. This puts in evidence an important improvement given by the aggregation of more PV subsystems. Instead, by considering the 1-min resolution, the results are around 40% for a single subsystem and 20% for the entire PV plant. They are quite different from the results obtained on

the Rovigo plant because of shorter observation period. If observed over a longer period, the figures would probably be more like those of the Rovigo plant.

Other studies were carried out in Spain in [7] and [8], they analysed the impacts on variability from increasing power plant size and from increasing geographical dispersion of up to eight Spanish PV systems in the capacity range of 48 kW_p to 9.5 MW_p, over distances of up to 60 km. The analysis is more detailed than most of the previous ones and covers time resolutions from 1 second up to 10 minutes. The studies show both the variability of differently sized individual systems and of different aggregates of systems on variability. It can be observed that in [7] the maximum 1-second step change is around 55% of capacity for the smallest system (48 kW_p) and merely around 5% for the largest system (9.5 MW_p), from one-year observations. The authors also report the 90th percentile of the step changes, plotted in Fig. 5 against the plant area. As can be seen, this quantity is considerably lower on shorter timescales and so there is a stronger smoothing effect, for example, considering a 100.000 m^2 power plant, the 90th percentile of the step changes distribution with a 1-second timescale is about 5% of the capacity, while with a 20-second timescale is 25%. In addition to that, it was noted that by combining six systems within distances of 60 km further reduces the total variability.

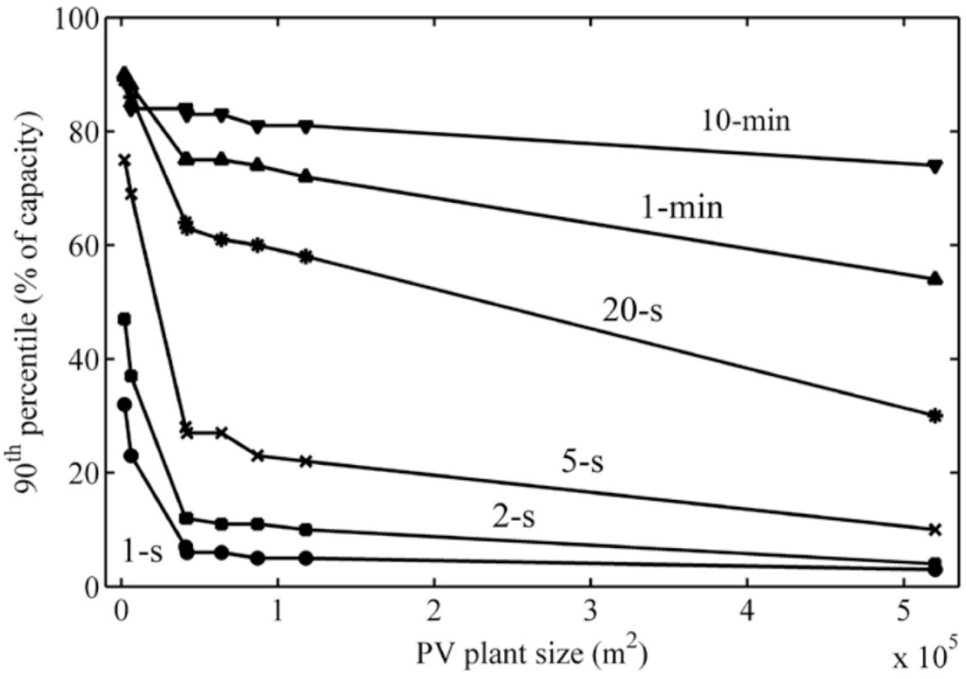

Fig. 5: 90th percentile of step changes in the output from differently sized PV plants based on values reported in [7].

In [8] the 99th percentiles are reported for all possible combinations of up to six plants. These quantities decrease from around 10% of capacity for one system to around 2% for six systems at the 1-s resolution, whereas by considering the 1-min resolution the drop is from 85% to 35%.

In addition, they use these data to find empirical relations between the number of dispersed PV plants grouped, the 99th percentile of the step change, and the time resolution. The result was that the percentile decreases systematically with the dispersion and increases with the time resolution. The decrease is from around 10% of capacity to 2% for the 1-s resolution, as mentioned, and from 90% to 50% for the 10-min resolution. In an attempt to generalize their findings, the authors also fit their data to an empirical function relating the 99th percentile of the step change to the PV plant size and the number of dispersed systems.

The smoothing effect has been also studied considering the dispersion over nation-wide distances. In [9] the electricity production data from 52 sites distributed over Japan were analysed. They investigated how the correlations of output fluctuations depends on the distances between the locations of PV systems, in conclusion, they found that over 1-min, sites more than about 50-100 km apart were uncorrelated and thus that there was a limit reached whereby adding more PV sites had no effect on reducing variability, since the variability introduced by the diurnal cycle eventually becomes larger than the cloud-induced variability. In [10] it was studied the electricity production from a power plant composed by 100 PV systems (with a size from 1 to 5 kW_p) spread over Germany, all monitored over one year with a 5-min resolution. It was observed that the power fluctuations were not higher than 5% of total capacity. In addition, it was found that the output power of the ensemble was never above 65% of the total capacity, so over a very large power plant the maximum achievable output power is considerable lower than the nominal capacity of the ensemble.

2.2 Observed variability and smoothing in solar radiation data

Up to now, many observations on the power output variability have been shown. It has to be considered that such power fluctuations depend on the irradiance variability.

Even though they are not directly proportional, the statistics about the smoothing due to the geographical dispersion provide very similar results.

One of the first study was made in [11] in 1997 on a Japanese system made of 9 sites in a 4 x 4 km grid, that is the size of a typical urban distribution network, using irradiance data with a 1-min resolution. The authors used a metric called the "*fluctuation factor*" fl to quantify the irradiance fluctuation. Such factor is obtained by the Root Mean Square value of the *fluctuant irradiance* $F(t)$, as equation (2) shows, where N is the number of irradiance samples.

$$fl = \sqrt{\frac{1}{N}\sum_{t=1}^{N} F(t)} \tag{2}$$

The *fluctuant irradiance* $F(t)$ is calculated by eliminating the *base component* from the irradiance data $G(t)$. Such *base component* is defined as the moving average of irradiance data G, the calculation of $F(t)$ is defined in equation (3) where m was set equal to 13 in order to analyse only the fluctuation shorter than 60 minutes

$$F(t) = G(t) - \frac{1}{2m+1}\sum_{\tau=-m}^{m} G(t+\tau) \tag{3}$$

The *fluctuation factor* fl was calculated for each location and for the averaged irradiance of the entire network. The authors did this calculation over the month of October and showed an average decrease of 40% of this factor. Even though there is a reduction of variability, it is not so evident because of short distances between the sites. A more stable power output can be obtained under no correlative irradiance fields in which individual irradiance fluctuations are geographically random.

The same network was used by [12] to study the impact of different typical weather types on fluctuations and smoothing across the grid, showing that on cloudy days there is a stronger smoothing effect because of the irradiance fluctuations, while clear and rainy days are scarcely influenced by the short-time moving clouds.

In a more recent study, [13] used measurements of global radiation on the horizontal plane at four sites across the State of Colorado, located between 19 and 197 km apart, to study the smoothing effect. The maximum ramp rate, using data covering one whole year with a 5-minute resolution, was between 161 and 189 $Wm^{-2}min^{-1}$ for

individual sites and 112 $Wm^{-2}min^{-1}$ for the four sites combined. It highlights that there is a reduction of variability in irradiance data too.

By considering that ramp rates in Wm^{-2} depend both on the sun's deterministic position in the sky and on cloud movements, it is better to analyse the clearness index rather than the absolute irradiance. The clearness index is defined as the ratio between the actual radiation and the clear-sky radiation, where the clear-sky radiation is the irradiance time-series obtained in absence of clouds (in clear-sky condition), such index provides only the irradiance components due to moving clouds.

In [14], the authors used the clearness index determined from 1-min global horizontal radiation data covering one whole year to quantify the smoothing from dispersion of 23 PV sites over distances from 20 to 440 km. The authors found that the 99.7^{th} percentile (of the ramp rates distributions) dropped from 0.58 for one site to 0.19 for five sites, and to 0.09 for all 23 sites. Since the clearness index and the PV output power are related, it is not surprising that these results are comparable in size to the 99.7^{th} percentile values for the PV output in [4] mentioned above. In [15], it was used a one-month series of 1-second data for the clearness index at six sites, separated by less than 3 km, in San Diego, to determine fluctuations on different time scales with a spectral analysis. The variability was quantified by using two metrics; the *fluctuation power index* (*fpi*) and the *variability reduction* (*VR*). The *fpi* is an index which represents the average power fluctuation at each timescale [16]. The *VR* is defined as the *fpi* for one site divided by the *fpi* for a set of aggregated sites. It ranges between 1 and the number of sites, respectively when the sites are dependent and independent on the considered timescale. A higher *VR* means a larger reduction in fluctuations, while a *VR* equal to 1 means no reduction in variability. Thus, it shows the reduction in variability with aggregation. One of the results from the study is that the *VR* is around 6 (the number of sites) for short timescales (shorter than 4-min) and approaches 1 for longer timescales (about 1-hour). This points out that there is not reduction of variability when the timescale is quite long because the locations are strongly correlated. A similar effect can be seen in Fig. 5 above where we observe that the variability is practically independent by the plant size for longer times scales.

These studies, among others, confirm that the variability associated with a fleet of PV systems is significantly lower than the variability associated with a single PV system [3]. Therefore, by observing the variability of the irradiance data, it can be noted that the conclusions are very similar to the one deduced by the studies on the power data.

2.3 PV Output Variability Models

Up to now, many data analysis have been shown to understand the effect of dispersion of the power plant on the variability reduction. In order to help the PV planners to design fleet distribution and sizing storage systems, many methods have been shown in literature to estimate changes in output power with high reliability.

In the subparagraph 2.3.1, it is illustrated a method to estimate the output power of a power plant composed by a single PV system, by knowing only a single irradiance time-series and the extension of the plant, so the smoothing effect due to the plant size is modelled.

In the subparagraphs 2.3.2 and 2.3.3, two methods to estimate the output variability of a generic distributed power plant are presented, such models have strong mathematical foundation. The second one is particularly interesting because it allows to calculate the reduction of variability that results by aggregation of more PV systems only by knowing the dispersion of the plant over a region and the cloud speed.

2.3.1 Smoothing of power fluctuations in a centralized PV plant

Numerous studies, using data with a time resolution of seconds or minutes, analysed power fluctuations in the frequency domain and the conclusion has been that large PV plant causes bigger attenuation of high frequencies than small plants.

An important study has been done in Spain analysing, for one year and with a time resolution of 1 second, the irradiance signal and the power output of six PV plants. The power of the plants ranges from 1 to 9.5 MW and the plants are scattered over about 1000 km^2 [17].

In Fig. 6 we can see the irradiance G_N and the output power P_N recorded at the Milagro PV plant (with a peak power of 9.5 MW). It can be noted that the output power has

the same general behaviour of the irradiance time-series, but it is a smoothed version of the irradiance.

The big size of the PV plant in comparison with the discrete character of the irradiance sensor causes a smoother power curve.

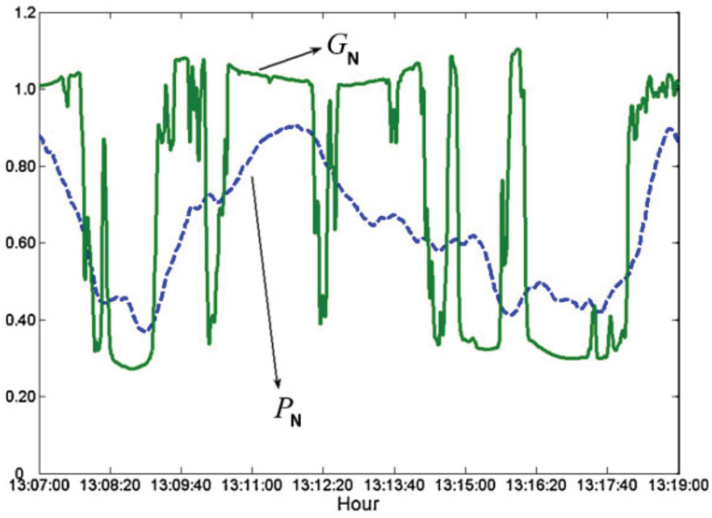

Fig. 6: Irradiance, G_N, and output power, P_N, both normalized to their nominal values, recorded at Milagro site during a 12-min period on 12 August [17].

The periodicity of the solar resource enables a Fourier analysis, and using a Fast Fourier Transform (FFT) algorithm, the Discrete Fourier Transform (DFT) of the irradiance signal G_N (recorded along a full year) was calculated. The same was done for the normalized PV plant power signals. In Fig. 7 we can see the power spectra calculated for the Milagro and the Sesma plant as well as the spectra of the irradiance G_N. The power spectra can be approximately described by two functions of the form $f^{-0.7}$ and $f^{-1.7}$ (the superscript indicates the slope coefficient), the cross point between the two regions defines the cut-off frequency f_c.

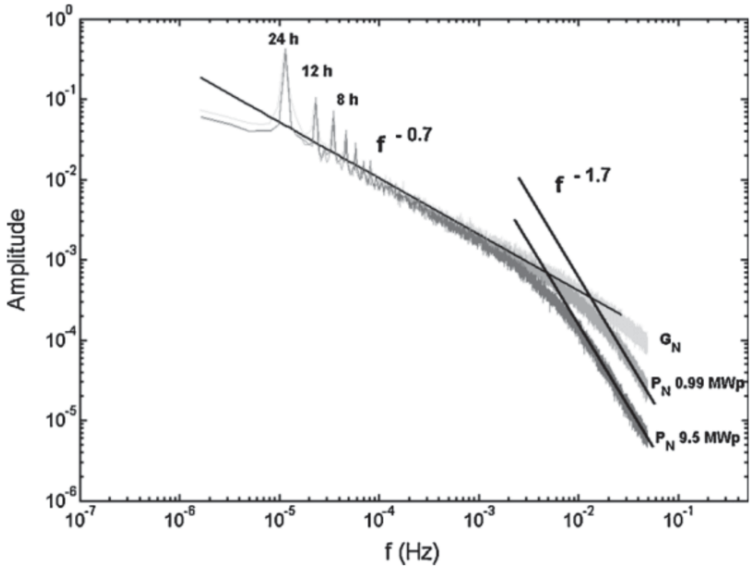

Fig. 7: Spectrum of the irradiance G_N recorded at Milagro (the linear region can be well fitted by a function of the form $f^{0.7}$), and the spectrum of the Output Power P_N at Sesma and Milagro during 1 year (here there are two linear region, they can be fitted by functions of the form $f^{0.7}$ and $f^{1.7}$) [17].

In other words, regarding power fluctuations, the PV plant size can be interpreted as a first order low-pass filter for the irradiance signal. The same analysis was done for all the PV plants (with different areas and power sizes) and so different values of the cut-off frequencies were found. As expected, the bigger the PV plant, the lower the cut-off frequency.

Table 3: Characteristics and cut-off frequency, f_c, estimated for each PV plant used in [17].

PV plants	Nominal power (kW)	Area S (Ha)	Cut-off frequency (Hz)
Milagro Section 2	100	0.63	0.026
Arguedas	775	4.1	0.0098
Sesma	800	4.2	0.0088
Cintruénigo	1155	6.4	0.0081
Rada	1400	8.7	0.0072
Castejon	2000	11.8	0.0069
Milagro	7243	52	0.0032

The relation between f_c and the PV plant area S is shown in Fig. 8. At large frequencies it is well fitted by the equation (4)

$$f_c = a \cdot S^b \tag{4}$$

where $a = 0.0204$ Hz/Ha and $b = -0.499$. So, the smoothing effect is practically given by a $1/\sqrt{S}$ law.

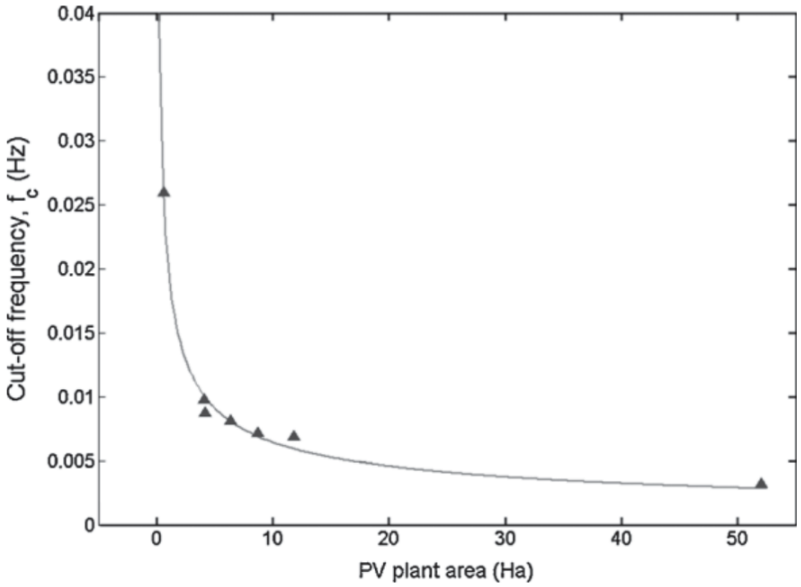

Fig. 8: Cut-off frequency as a function of the PV plant area S [17].

In conclusion, for a given irradiance time series $G(t)$, we can obtain the output power $P_s(t)$ using the transfer function $H(s)$ in the Laplace domain defined in (5), that models the PV plant as a first order filter with a pole $\tau = (2\pi f_c)^{-1}$, which depends on the PV plant area S.

$$H(s) = \frac{\mathcal{L}\{P_s(t)\}}{\mathcal{L}\{G(t)\}} = \frac{K}{\tau s + 1} = \frac{K}{\left(\frac{\sqrt{S}}{2\pi \cdot 0.02}\right)s + 1} \tag{5}$$

where s is the Laplace variable, $G(t)$ is the time-series irradiance data, K relates the power and the irradiance.

If we want to obtain the equivalent smoothed irradiance $G_s(t)$, we can use the transfer function (6), where K is equal to 1.

$$H(s) = \frac{\mathcal{L}\{G_s(t)\}}{\mathcal{L}\{G(t)\}} = \frac{1}{\tau s + 1} = \frac{1}{\left(\frac{\sqrt{S}}{2\pi \cdot 0.02}\right)s + 1} \tag{6}$$

By using the equation (5), it is possible simulating the output power and then compare it with the real data in order to validate the model. This has been done at the Milagro (52 Ha) site, analysing in the frequency domain the error between real and simulated output power, this DFT error is shown in Fig. 9 and it can be observed that it is below 2 % for all the frequency range.

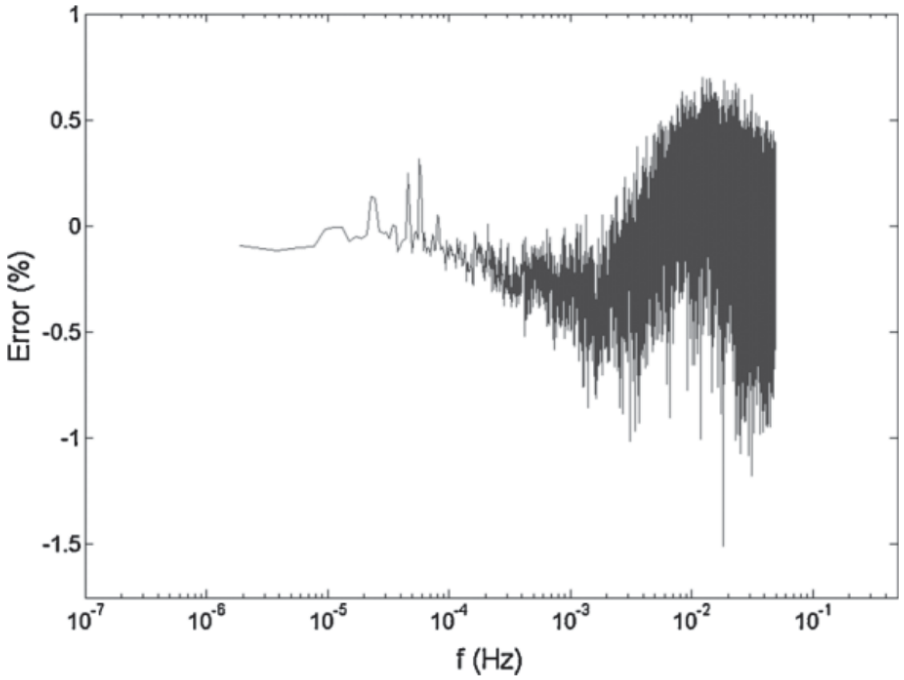

Fig. 9: DFT error between real and simulated output power at Milagro for 1-year data [17].

Another check has been done on the maximum short-term (below 10 minutes) daily power fluctuations $\Delta P_{\Delta t,max}$, this quantity is of particular interest from the grid operator viewpoint [7]. Therefore the daily maximum power fluctuation error $E_{\Delta t,d}$ is defined in (7) as the difference between the real daily maximum power fluctuation and the simulated one.

$$E_{\Delta t,d}(\%) = \left[\Delta P_{\Delta t,max,real} - \Delta P_{\Delta t,max,sim}\right] \quad (7)$$

Then, this error was calculated for each day of the analysed year at different timescales (from 1 second to 600 seconds), and the annual evolution statistical analysis is shown in Fig. 10.

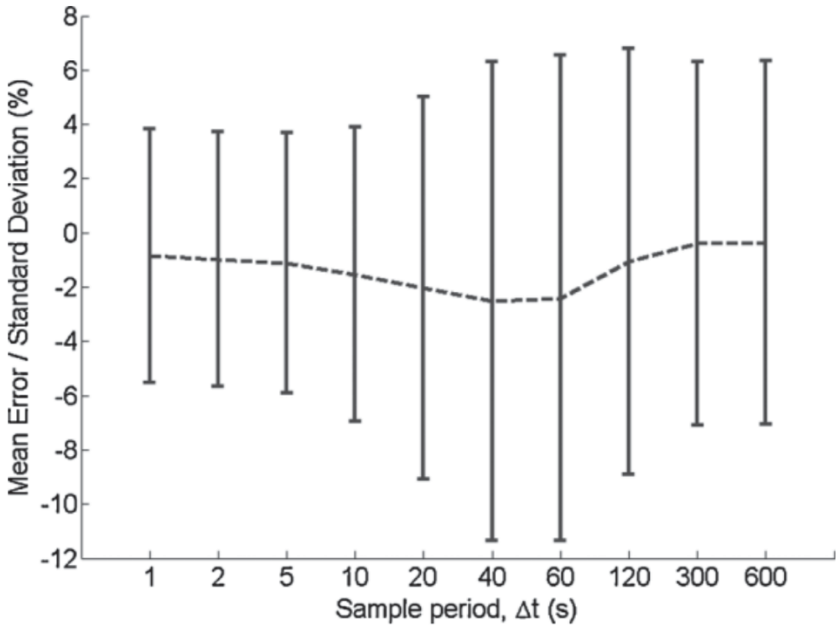

Fig. 10: Mean error (dashed line) and standard deviation obtained for the annual evolution of the daily maximum power fluctuation error $E_{\Delta t,d}(\%)$ considering 365 days for each Δt [17].

It can be easily observed that the mean error does not exceed 2% over all considered timescales, so it can be concluded that the developed model is precise and reliable.

2.3.2 Wavelet Variability Model (WVM)

In [18] it was proposed a model for simulating PV power plant output given a single irradiance point sensor time-series, knowledge of the power plant footprint, PV density (Watts of installed capacity per square meter) and a correlation scaling coefficient by determining the geographic smoothing that will occur over the area of the plant. As hypothesis, the simulated power plant may have any density of PV coverage: it may be distributed generation with low PV density, centrally located with high PV density, or any combination of both.

By using this method, the correlation between PV systems is assumed depending only on distance, so the correlation is not influenced by the direction of the moving clouds.
The main steps to obtain simulated power plant output are described in the following part:

a) The first step is to normalize irradiance time-series in order to obtain a stationary signal, to do this operation Solis model [19] (which models the

irradiance time-series obtained without clouds) can be used. The resulting signal is then decomposed into its components at various timescale by using a wavelet transform (it is a mathematical tool to analyse a signal in both temporal and frequency domain), the output are the different components (or "modes") $\omega_{\bar{t}}(t)$ showing fluctuations at these various timescales.

b) The power plant is discretized into different locations. For distributed plants, a single location is one rooftop PV system. For centralized plants, a single location is divided into groups of PV modules, and each group is considered as a location.

c) The correlations $\rho(d_{m,n}, \bar{t})$ between locations within the power plant is given by the equation (8) where it can be observed that such correlations depends on $d_{m,n}$ and $\bar{t}$. $d_{m,n}$ is the distance between locations m and n, $\bar{t}$ is the timescale of fluctuations and A is a correlation scaling factor.

$$\rho\left(d_{m,n}, \bar{t}\right) = e^{\left(-\frac{1}{A}\frac{d_{m,n}}{\bar{t}}\right)} \quad (8)$$

The A value can be found using a small network of irradiance sensors where the correlations and distances are known in advance and the previous equation can be used inversely. If the sensors network is not available, another way of computing the A value is proposed in [20] and reported in (9), it is merely proportional to cloud speed cs expressed in *m/s*

$$A = 0.5 \times cs \quad (9)$$

Smaller A values result in lower correlations between locations.

From the correlations, the "variability reduction" $VR(\bar{t})$ (already defined in the paragraph 2.2) can be calculate as shown by equation (10) where N is the total number of locations.

$$VR(\bar{t}) = \frac{N^2}{\sum_{m=1}^{N}\sum_{n=1}^{N}\rho\left(d_{m,n}, \bar{t}\right)} \quad (10)$$

d) By combining the wavelet modes $\omega_{\bar{t}}(t)$ with $VR(\bar{t})$, at the respective timescale, it is possible to simulate the normalized plant average irradiance. These quantities are related by equation (11).

$$\omega_{\bar{t}}^{sim}(t) = \frac{\omega_{\bar{t}}(t)}{\sqrt{VR(\bar{t})}} \quad (11)$$

At this point, the resulting modes $\omega_{\tilde{t}}^{sim}(t)$ at the different timescales are summed to create a simulated normalized irradiance, that is spatially averaged over the power plant.

e) Simulated output power is obtained by multiplying the simulated irradiance by a clear-sky model for output power (it takes also the PV density into account), such operation is the opposite one of the normalization done in a).

Although the fluctuations of the simulated power output and actual power plant output are not expected to match perfectly since that only a single irradiance sensor is used as input, the statistics of the fluctuations are expected to agree.

The model was validated at the 2 MW distributed residential rooftop plant in Ota City (Japan) and the 48 MW central power plant in Copper Mountain (USA). For both locations, Fluctuation Power Index (presented in paragraph 2.2) was compared to the actual one and it resulted that they matched, also the Probability Density Functions (PDF) for 1, 10, 30, 60 seconds were compared and the error was found being pretty negligible.

In conclusion, the main benefit of using the WVM is the reasonable input requirements to estimate the power output fluctuations, and at the same time the results are very precise.

2.3.3 Dispersion Factor Method

In order to quantify the power output variability from a fleet of PV systems, an innovative method was presented by Hoff and Perez [21]. This method allows us to quantify the variability taking in consideration the scenario in which there is an ensemble of equally-spaced, identical PV systems. This situation can help in the analysis of different scenarios from a single central station (N closely-spaced installations) to a set of distributed PV systems (N widely-spaced installations).

The output power variability is defined in (12) as a measure of the PV fleet's power output changes over a selected sampling Time Interval and analysis period relative to PV fleet capacity

$$\sigma_{\Delta t}^{Fleet} = \frac{1}{C}\sqrt{Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{P}_{\Delta t}^{n}\right]} \quad (12)$$

where C denotes the total installed nominal capacity of the fleet, Var identifies the variance and $\Delta \boldsymbol{P}_{\Delta t}^{n}$ is the time-series of changes in power at the n^{th} PV installation using a sampling Time Interval of Δt. Since that the power production can be considerate approximately directly proportional to the Irradiance, the model solution (presented later in Table 4) is calculated and validated just considering the irradiance time-series.

The central variable of the proposed model is the Relative Output Variability, here it is denoted with ROV. This quantity is defined in (13) as the ratio between the Output Power Variability of a PV plant and the Output Power Variability of a single PV system of the power plant. It ranges between 0 and 1 and quantifies the reduction of variability associated with the geographical dispersion of the power plant over a region.

$$ROV = \frac{\sigma_{\Delta t}^{Distributed}}{\sigma_{\Delta t}^{system}} \quad (13)$$

To better understand, a value of this ratio equal to 0.5 means that the variability of the entire plant is the half of the variability of the single station. Obviously, as ROV is lower, the fluctuations are more reduced.

To capture the relationship among PV plant configuration (the number of systems and the distance among the systems), Cloud Speed and the Time Interval over we consider the variability, another variable was defined, that is called Dispersion Factor (indicated by D) and it is dimensionless, its physical meaning is the number of Time Intervals required for a cloud disturbance to pass across the entire PV plant. So, the Dispersion Factor is defined in (14) where L is the length of the PV plant in the direction of the clouds motion, v is the Cloud Speed, and Δt is the considered Time Interval.

$$D = \frac{L}{v \cdot \Delta t} \quad (14)$$

We can notice that D increases when the cloud speed decreases, and when the distance between PV systems increases (because the cloud disturbance takes more Time Intervals to cover all PV systems).

Fig. 11 shows three scenarios in which the cloud speed changes, so we have different Dispersion Factors. The fast-moving cloud in the top section of the figure crosses the PV plant in time $2\Delta t$, and thus D equals 2. The medium transit speed requires time $4\Delta t$ for a cloud to pass over the PV plant, and thus D equals 4. In the last case, the slowest one, we have a Dispersion Factor of 8.

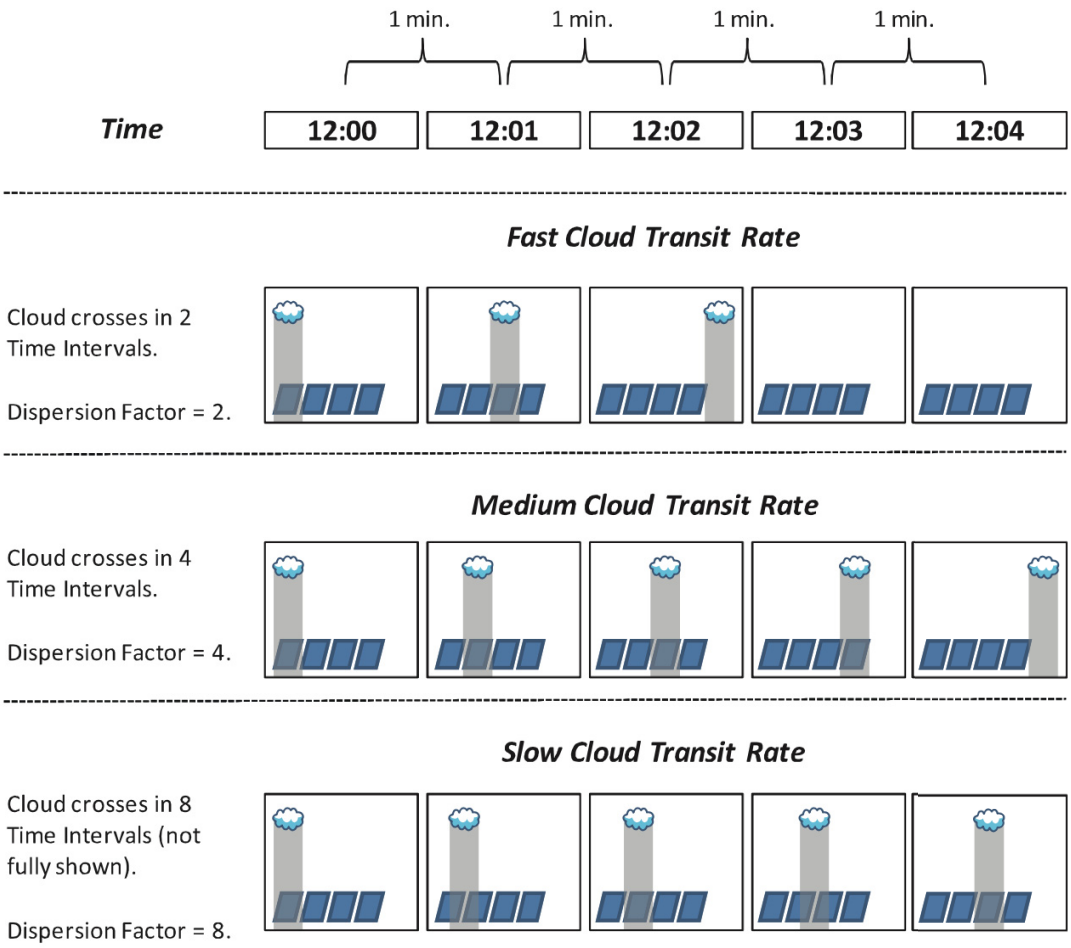

Fig. 11: Three scenarios with 4 PV systems and Time Interval is equal to 1 min but with different Cloud Transit Rates [21].

In order to calculate the Relative Output Variability, the proposed model provides four different solutions for the Output Variability $\sigma_{\Delta t}^{Fleet}$, depending on D value. Therefore, it can be possible to define four "regions" for D value:

- Crowded Region: The number N of PV systems is greater than Dispersion Factor. We can observe this situation in the top section of Fig. 11, where we see that, in one Time Interval, the cloud disturbance shadows two PV systems.

- Optimal Point: In this situation, the number N of PV systems equals the Dispersion Factor. This is shown in the middle section of Fig. 11, each PV system is affected by the cloud disturbance in exactly one Time Interval.
- Limited Region: The number N of PV systems is less than the Dispersion Factor. We see that in the bottom section of Fig. 11, so a single Time Interval is not enough for a cloud disturbance for reaching the next PV system.
- Spacious Region: This is an extension of the previous region in which the short-term fluctuations of each PV system become independent of each other, so the number N of PV systems is much less than the Dispersion Factor.

The authors calculated mathematically the expected Output Variability in each region, the results are shown in Table 4, where $\sigma^1_{D\Delta t}$ is the variability of a single PV system evaluated over a Time Interval equal to $D\Delta t$, $\sigma^1_{N\Delta t}$ is the same but evaluated over a Time Interval equal to $N\Delta t$ and $\sigma^1_{\Delta t}$ is evaluated over Δt.

Table 4: Output Variability solution of the proposed model [21].

Crowed Region	**Optimal Point**	**Limited Region**	**Spacious Region**
$\sigma^{Fleet}_{\Delta t} = \frac{\sigma^1_{D\Delta t}}{D}$	$\sigma^{Fleet}_{\Delta t} = \frac{\sigma^1_{N\Delta t}}{N}$	$\sigma^{Fleet}_{\Delta t} < \frac{\sigma^1_{\Delta t}}{\sqrt{N}}$	$\sigma^{Fleet}_{\Delta t} = \frac{\sigma^1_{\Delta t}}{\sqrt{N}}$

So, the plot of the resulting model is showed in Fig. 12 for N PV systems. It illustrates the Relative Output Variability as a function of the Dispersion Factor, in the first region the Relative Output Variability goes down and it reaches a minimum when D is equal to N, then it increases in the Limited Region and then it reaches a value of $1/\sqrt{N}$. Moreover, it can be noted in Spacious Region, that by increasing the distance between the systems, for a fixed number of locations, does not minimize Relative Output Variability, because the power outputs are already uncorrelated.

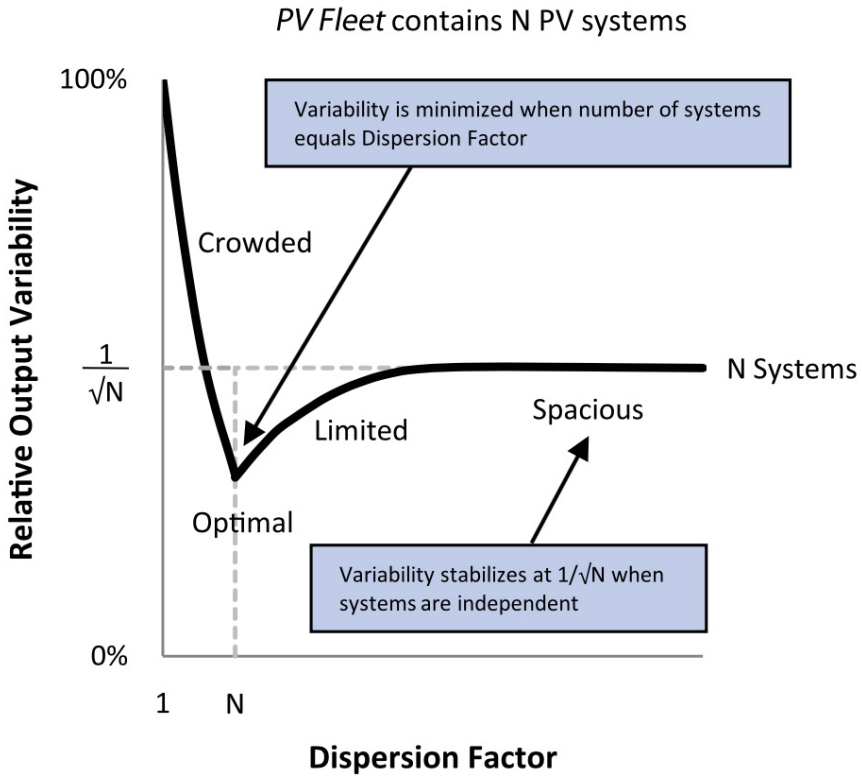

Fig. 12: Relative Output Variability as a function of the Dispersion Factor [21].

However, this relation depends also on the number N of PV systems which constitute the power plant. Fig. 13 shows what happens to the curve when the number of systems increases to $4N$. It is worth noting that when we quadruple the number of systems, in the spacious region we have the half of the Relative Output Variability. So as the number of systems increases, the Relative Output Variability decreases.

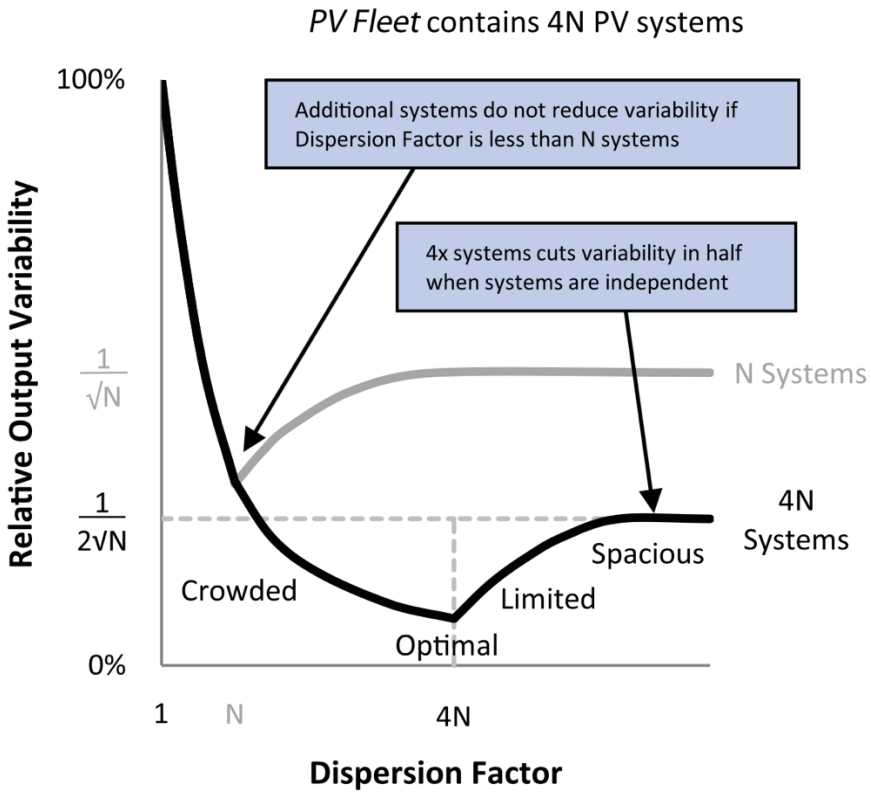

Fig. 13: Differences between the case with N systems and that one with 4N systems [21].

By simulating many different power plants with different geographical extensions, number of systems, cloud transit speed, and Time Interval, but using real irradiance data, this method was validated and it was found that the mean Relative Output Variability has a $1/\sqrt{N}$ behaviour when the systems are uncorrelated (Dispersion Factor is in Spacious Region) as shown in Fig. 14. It means that it is possible to estimate the reduction of variability from one single location to the entire power plant only by knowing the number N of system that composes the plant.

A common behaviour was observed in the Crowded Region (when D is smaller than N), as shown in Fig. 15. It points out that the reduction of variability depends only on D value when the systems are very close to each other.

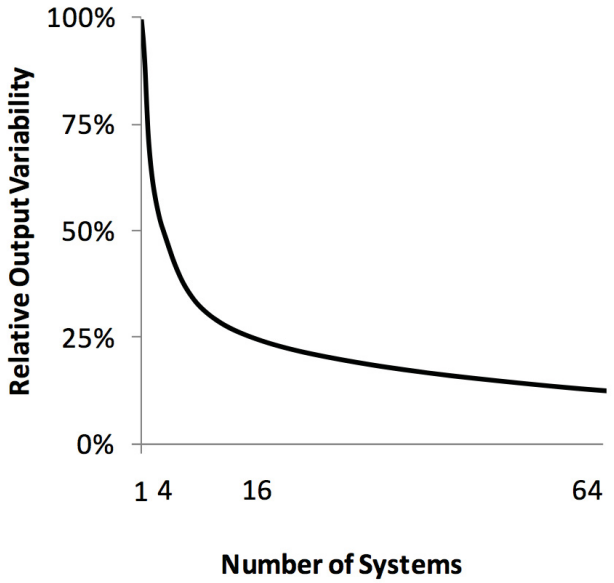

Fig. 14: Relative Output Variability for Spacious Region [21].

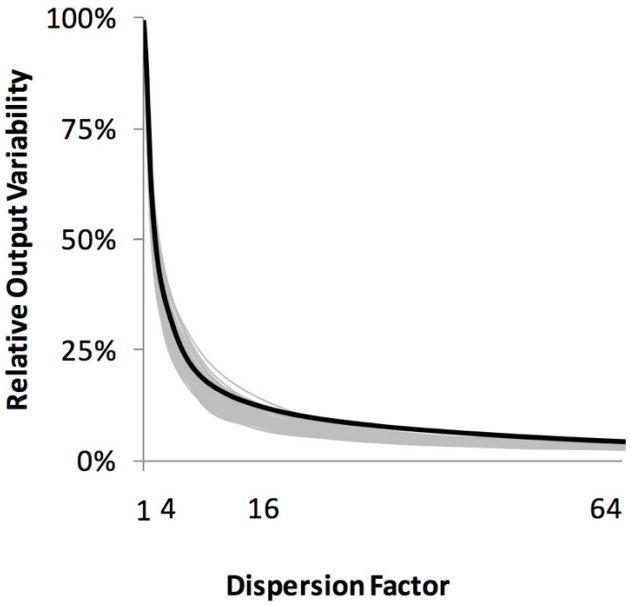

Fig. 15: Relative Output Variability for Crowded Region [21].

Other authors tried to test the validity of this model. In [7], an empirical model to estimate the 90th percentile of the power fluctuations (as a function of PV plant area) was proposed by fitting the monitored data (see Fig. 5). The model was built considering centralized power plant (so each PV system can be considered very close to each other) and it is represented by equation (15), where S is the plant area, *m* is some proportional constant and *c* represents the attenuation of power fluctuations because of the plant area S, such values are different for each time interval.

$$90^{th}(\Delta P_{\Delta t}) = mS^{-c} \quad (15)$$

By analysing data with short Time Interval like 1-second or 2-second, *c* is equal to 0.5 so the percentile decreases proportionally to $1/\sqrt{S}$, where $\sqrt{S}$ can be seen as the extension of the plant in one direction, so as L. It agrees with the expression considered for the crowded region (reported in Table 4) where the output variability is inversely proportional to D (so inversely proportional to L, the PV plant length). Such correspondence can be clearly seen by observing the relation (16) where it can be possible to see how the variability (interpreted both 90th percentile and standard deviation) linearly decreases with L.

$$90^{th}(\Delta P_{\Delta t}) = \frac{m}{\sqrt{S}} = \frac{m}{L} \quad \Leftrightarrow \quad \sigma_{\Delta t}^{Fleet} = \frac{\sigma_{D\Delta t}^{1}}{D} = \frac{\sigma_{D\Delta t}^{1} v \Delta t}{L} \tag{16}$$

For lower resolution data (higher Time Interval Δt), *c* tends to zero, so the plant size S practically does not influence the fluctuations. This is in line with the result in Crowded Region because D, that is inversely proportional to Δt, has a very low value and the Relative Output Variability approaches one for lower D, it means that it is not influenced by S. This can be physically explained by considering that clouds have time to pass over the entire PV plant within a single Time Interval because it is very long.

Successively, another model was proposed in [8] considering the 99th percentile of the step change distribution for an arbitrary power plant composed by N sufficiently uncorrelated PV systems, the relation is reported in the equation (17) where S denotes the size of a single PV system and $\Delta P_{600,1}$ is the 10-min step change considering a single PV system.

$$99^{th}\left(\Delta P_{\Delta t,N}\right) = 99^{th}\left(\Delta P_{600,1}\right)(1 - e^{-0.24\Delta t})S^{-c}N^{-a} \tag{17}$$

To obtain the parameters *c* and *a,* the expression is fitted to data describing the smoothing due to both plant size and site aggregation. The *c* parameter keeps the same behaviour of the previous model. The *a* parameter is close to 0.5 for large time scale, this result agrees with the model in Spacious Region, because of the proportionality of the standard deviation to $1/\sqrt{N}$. In addition, always considering large time timescale, it was noted that the power fluctuations are much more attenuated by N than S.

The authors use the equation (17) to extrapolate the 99th percentile of step changes for a 100 MW_p PV capacity by considering in different scenarios. The results are shown in Table 5.

It is possible to compare the 1-min step changes to the results discussed previously in [5]. The value estimated for the centralized system is 70.9% of capacity, whereas in Rovigo plant was about 45%. The main reason is that the model developed in [8] overestimates the percentile for larger areas for longer timescales (like 1-min or 10-min) and obviously, also the local weather conditions influences the results [3].

Table 5: 99th percentile of $\Delta P_{\Delta t,N}$ values for different Δt and N, the total capacity of the PV plant is always the same [3].

N	S (Ha)	99th percentile (% of total capacity)		
		1-s	1-min	10-min
1	651	0.9	70.9	86.1
10	65.1	0.5	18.6	31.3
100	6.51	0.2	4.9	11.3
1000	0.651	0.1	1.3	4.1
10000	0.0651	0.1	0.3	1.5

2.4 Final considerations

The results discussed over this chapter are useful because they show the behaviour of the output variability by changing the footprint of the PV power plant. In particular, It is important to estimate such variability during the planning phase in order to choose how to split the power plant into different locations.

In the following part of this thesis work, the method shown in 2.3.1 is used to model the smoothing effect due to the physical extension of a single PV system, in this way it is possible to calculate the simulated output power in a single location.

Whereas the Dispersion Factor Method, illustrated in 2.3.3, affirms that, when the systems are sufficiently far away among them, the power outputs of each location are uncorrelated and the variability of the entire fleet is related with the variability of a single location with a $1/\sqrt{N}$ law.

Such result has been validated only by considering irradiance data from a single discrete location (and not the power data). It is interesting to compare the results obtained by considering the simulated output power. In other words, the smoothing

effect, due to the physical extension of each PV system, was not taken into consideration during the evaluation of the reduction of variability.

This is done in order to check if the results are different. Therefore, the Relative Output Variability is going to be evaluated using the simulated output power to investigate if the smoothing effect on each PV system has a relevant effect on the reduction of variability (from a single PV system to the entire fleet) in a distributed PV power plant.

3 Estimating Output Power

In this chapter, it will be shown how the calculation of the Output Power has been performed with the ultimate goal of analysing and comparing the variability of different footprint of PV power plant in the chapter 4. Two different cases need to be distinguished, depending on PV power plant footprint, as shown in Fig. 16. If the power plant is just composed by one PV system located in a single location (see Fig. 16a), the method developed in the paragraph 2.3.1 can be used to model the smoothing effect due to the physical extension of the system. Otherwise, if the power plant is composed by more PV subsystems (see Fig. 16b) placed in more locations, a different approach must be used. In particular, in this master thesis, a method is developed and validated to estimate the variability of a PV power plant just by using some meteorological data (like irradiance data and cloud speed) of the specific geographical region.

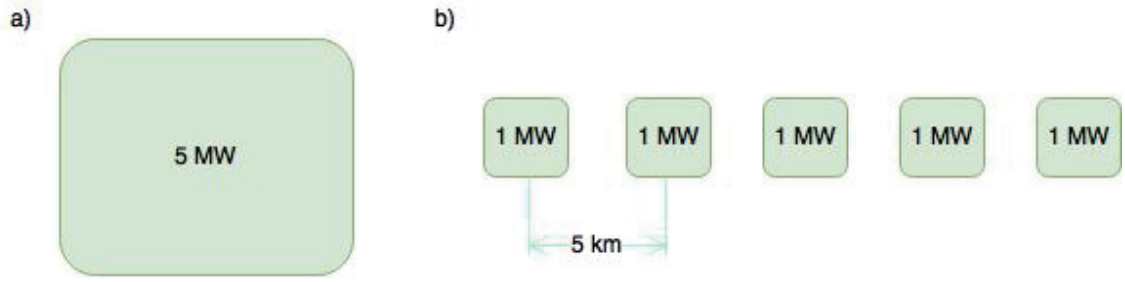

Fig. 16: a) A centralized power plant of 5 MW and b) a distributed PV fleet, in which the same capacity (5 MW) is divided on 5 subsystems, each one of 1 MW.

3.1 Dataset and assumption of PV systems

The dataset over which the methods shown here is applied, and used during the analysis performed in the chapter 4, is composed by the irradiance data of three months (May 2012, June 2012 and May 2013) from 8 a.m. to 6 p.m. captured by a single pyranometer CMP22 (Kipp&Zonen) located on TUT Solar PV Power Station Research Plant in Tampere (Finland) with a sampling frequency of 10 Hz, even if it was demonstrated in [22] that a sampling frequency of 1 Hz is high enough to identify all relevant transitions, so the irradiance data has been collected from the database with a frequency equal to 1 Hz.

The average speed of shadows passing over the power plant located there has been considered, in the present work, being equal to 13 m/s. Moreover, the main direction of moving clouds has been found being from south-west to north-east. The wind direction is useful because it is supposed that the subsystems of the distributed fleet are located along such direction. These data have been found in a previous study conducted in the Tampere region [23].

In addition, it was assumed a PV density of 153 kW/Ha, it has been calculated from the ratio of the capacity and the size of the typical power plant presented in [17]. This quantity relates the capacity of the system in kW with its physical size expressed in Hectare, for example, referring to Fig. 16a, a system with a capacity equals to 5 MW has an extension of 32.6 Ha. While, for one of the subsystems in Fig. 16b with a capacity of 1 MW, it is assumed an extension of 6.52 Ha (exactly one fifth of 32.6 Ha). Such number is useful because it is the main parameter of the transfer function (presented in 2.3.1) used to obtain the output power of a single system from the irradiance data.

While for a centralized power plant, the irradiance data measured by a single sensor are enough, for the power plant with a distributed footprint the irradiance data from each system of the fleet would be needed. Since only irradiance data of a single location is available in the Tampere region, a virtual PV fleet has been built in order to simulate the existence of the other subsystems.

3.2 Output Power Calculation for a Centralized Plant

In the case of a centralized system, the output power can be calculated as reported in the paragraph 2.3.1. The flowchart is reported in Fig. 17, just a low pass filtering has to be applied to obtain the simulated output power of the system. The input is the irradiance data time-series and the size of the system is used to calculate the cut-off frequency of the low-pass filter.

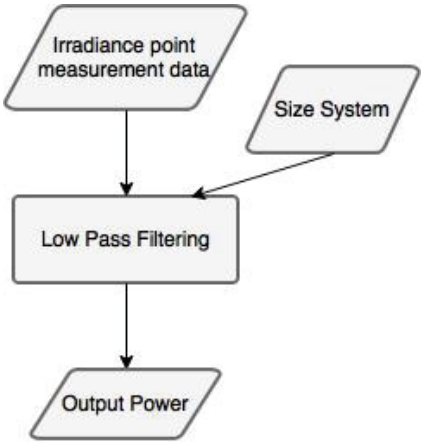

Fig. 17: Flowchart that describes how to calculate the output power for a centralized plant.

For illustrative purpose, Fig. 18 shows an irradiance data time-series captured during a 10-minutes period with a 1 second resolution in the afternoon on 2nd May 2012. The plot puts in evidence many irradiance transitions caused by edges of cloud shadows. The red-line in Fig. 19 displays the simulated output power calculated for a 5 MW centralized system (with an extension equal to 32.6 Ha) after the low-pass filtering with the corresponding cut-off frequency of 0.0035 Hz. It is shown a comparison with a proportional version (see the dashed line) of the irradiance data presented in Fig. 18, this line would represent the output power without considering the smoothing due to the size of the plant. It can be observed how the simulated power profile is smoother than the scaled version of the irradiance data.

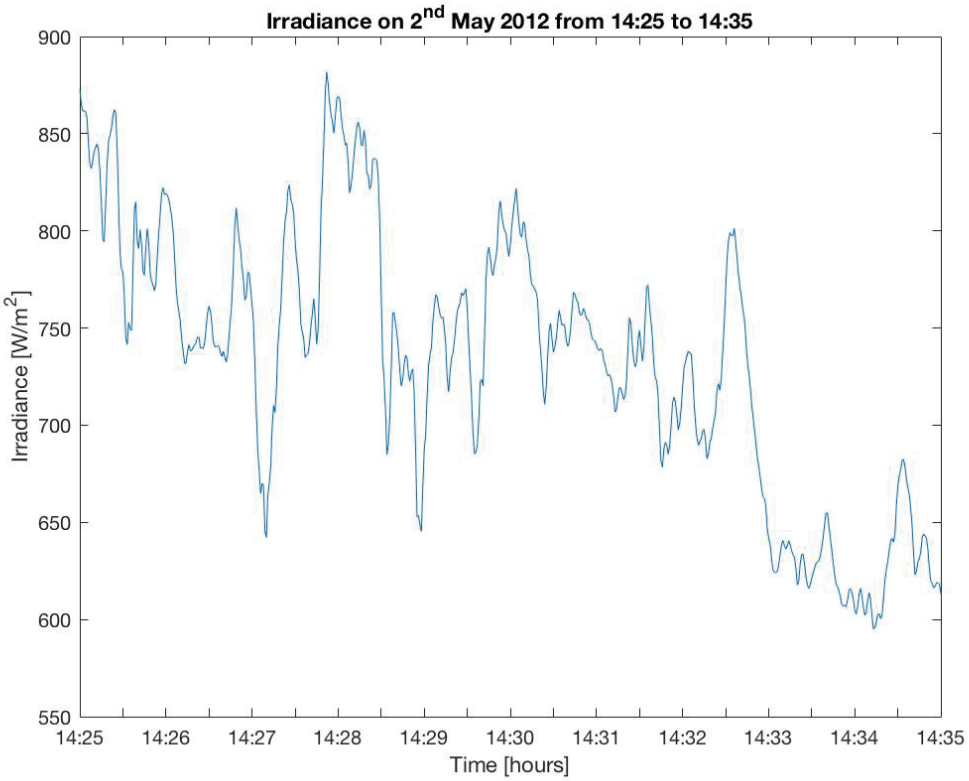

Fig. 18: Irradiance point measurement data on 2nd May 2012 from 14:25 to 14:35.

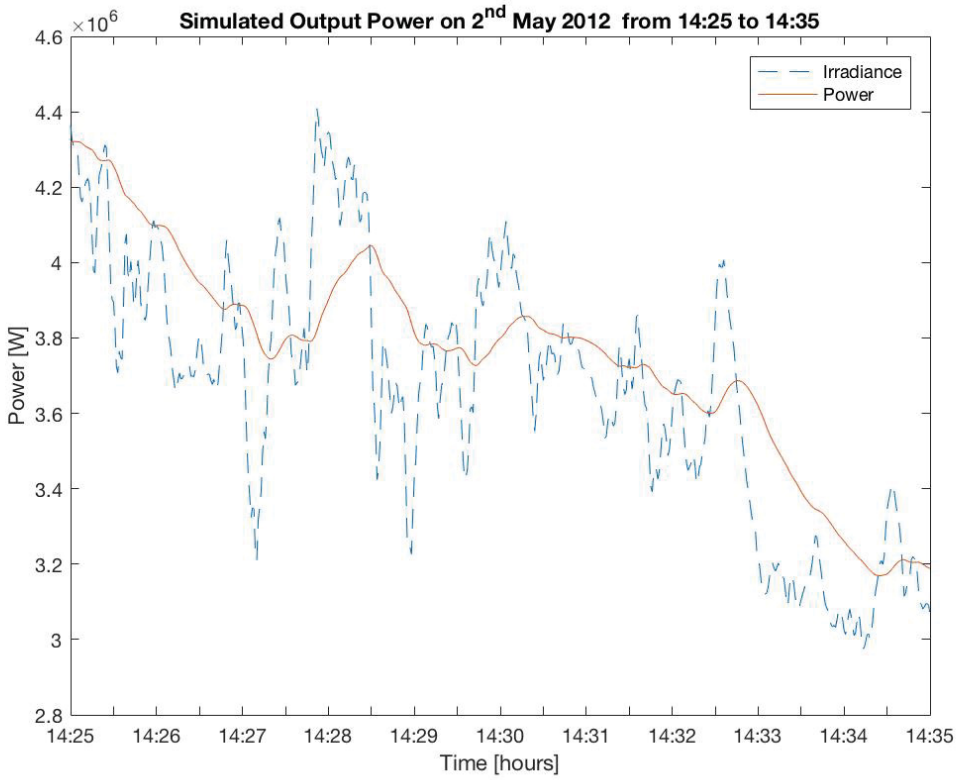

Fig. 19: Simulated output power on 2nd May 2012 from 14:25 to 14:35 and comparison with a scaled version of irradiance data.

3.3 Output Power Calculation for a Distributed Fleet

For the distributed system, because of lack of irradiance data originated by other pyranometers in the Tampere region, the irradiance point measurement data go through a different process, as shown in Fig. 20. The irradiance signal, measured at an actual location, is used by assuming that the cloud pattern moves at a constant speed across the virtual power plant. In this way, irradiance data for each virtual location are the same of the actual one, but shifted by knowing the average cloud speed and the distance between the PV systems, that is done in order to obtain the irradiance signals of N different locations distributed over the region. Then a low-pass filtering was applied to calculate the output power of each system and finally the various power outputs were summed.

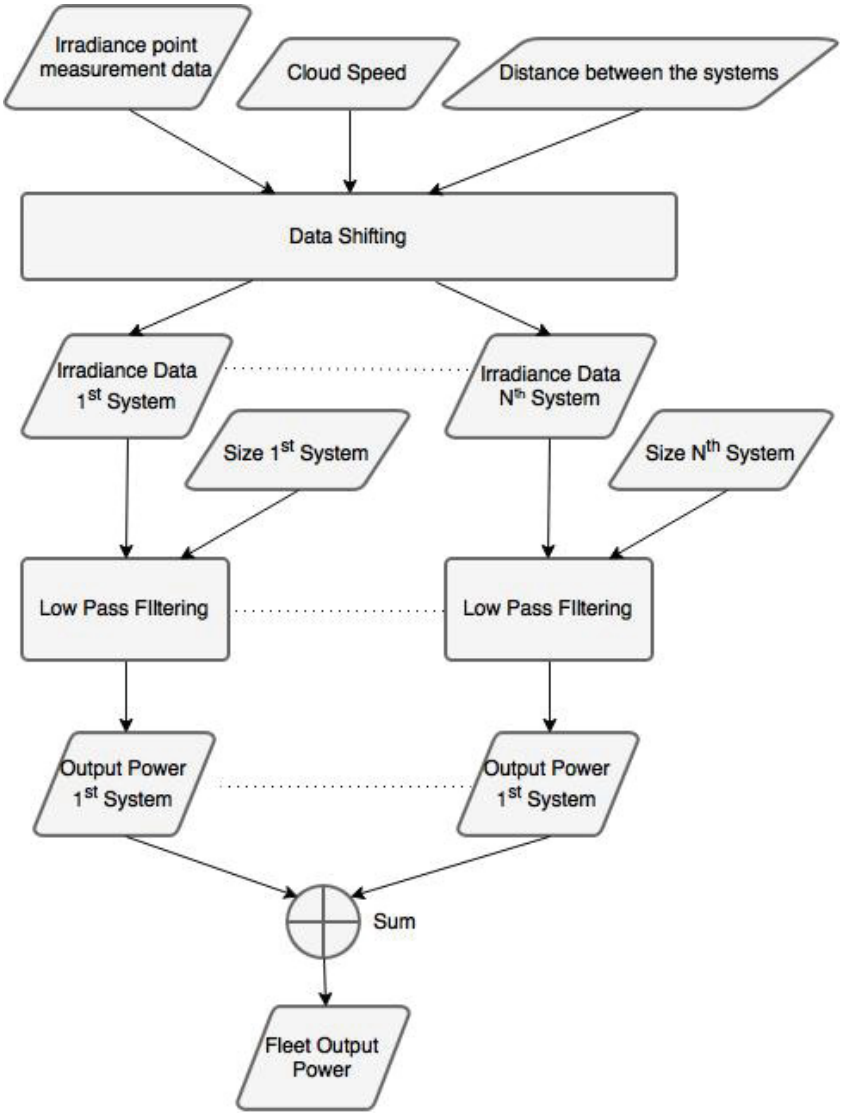

Fig. 20: Flowchart that describes how to obtain the fleet output power from irradiance point measurement data.

To better understand how the irradiance data time-series is used to calculate the irradiance data of another system, an example is presented considering two subsystems, each one with a supposed capacity of 1 MW. Let us remember that the average cloud speed v is 13 m/s and assume that the distance d between two adjacent system is 5 km. By considering these data and applying the note Velocity Formula from the Physics (see Eq. (18)), it can be found that a cloud takes about 6 minutes to move from a system to the next one, in this way the same irradiance data recorded by a single pyranometer can be assumed being the same of the virtual system, but just shifted, as shown in Fig. 21.

$$t = \frac{d}{v} = \frac{5000\ m}{13\ m/s} = 384\ s \approx 6\ min. \tag{18}$$

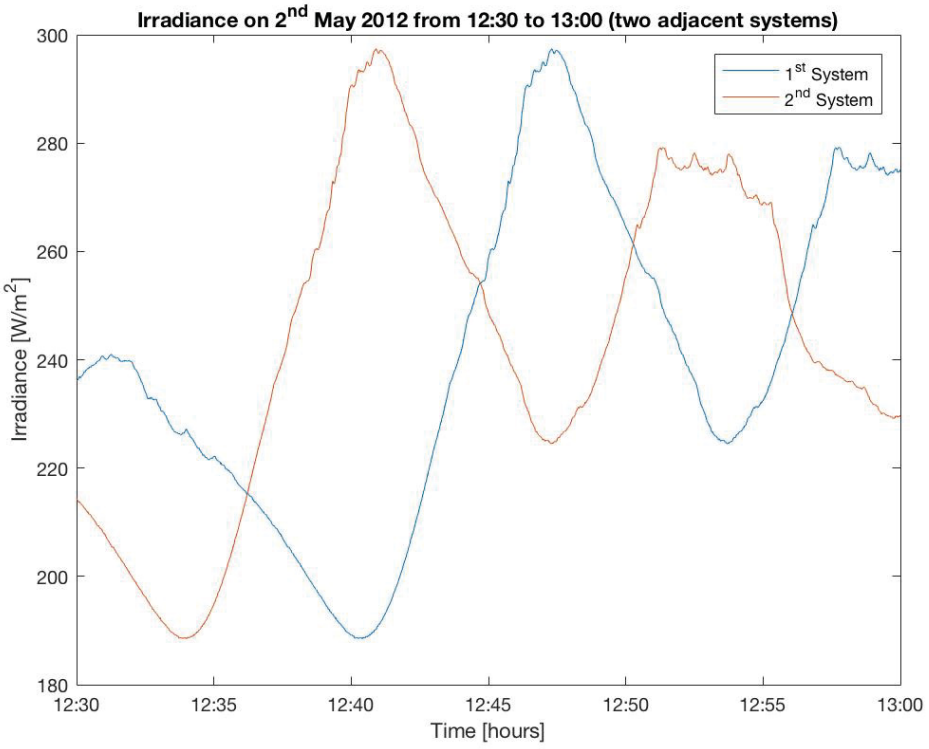

Fig. 21: Actual measurement data and its shifted version to simulate the existence of another system.

The obtained irradiance data time-series are used to calculate the output power of each system, that is done by applying a low pass filtering, and then all system output powers are summed to obtain the fleet output power. Fig. 22 displays (in blue) the simulated output power of a distributed fleet composed by 5 PV systems, each one has

a 1 MW capacity. It can be easily seen by sight that, compared to the output power of a centralized system (see the red line), the output power of a distributed system is more smoothed than the centralized solution, this is due to the geographical distribution of the fleet.

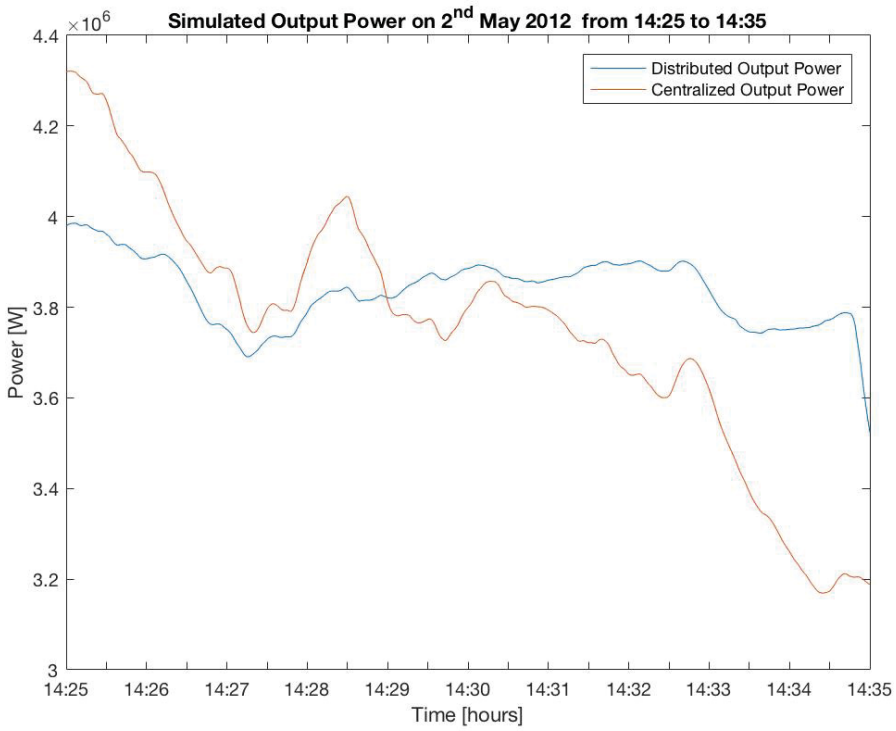

Fig. 22: Simulated output power on 2nd May 2012 from 14:25 to 14:35 for a distributed fleet and a centralized system.

3.4 Evaluation of the variability

Since that the aim of this work is to compare the output variability of different power plant layouts, it must be defined a way to calculate the output variability of a PV power plant.

This variability is represented by a random variable ΔP, defined in (19) that symbolizes the set of changes using a Time Interval of Δt, which represents the interval between two considered values of the output power. The variability, calculated in this way, is the normalized (to the plant capacity C) difference between two instantaneous output power values, and then multiplied by 100 to express this quantity in percentage of power plant capacity.

$$\Delta P(t) = \frac{[P(t + \Delta t) - P(t)]}{\mathrm{C}} \times 100 \tag{19}$$

An illustrative example of this quantity is shown in Fig. 23, where the output changes are calculated for a 5 MW power plant considering both solutions (centralized and distributed) with a Δt equal to 20 seconds. For this power plant, a fluctuation equal to 1% means that the difference between the current output power value produced and the value recorded 20 seconds before is 50 kW. This plot confirms that, in general, a distributed approach produces a less fluctuated output.

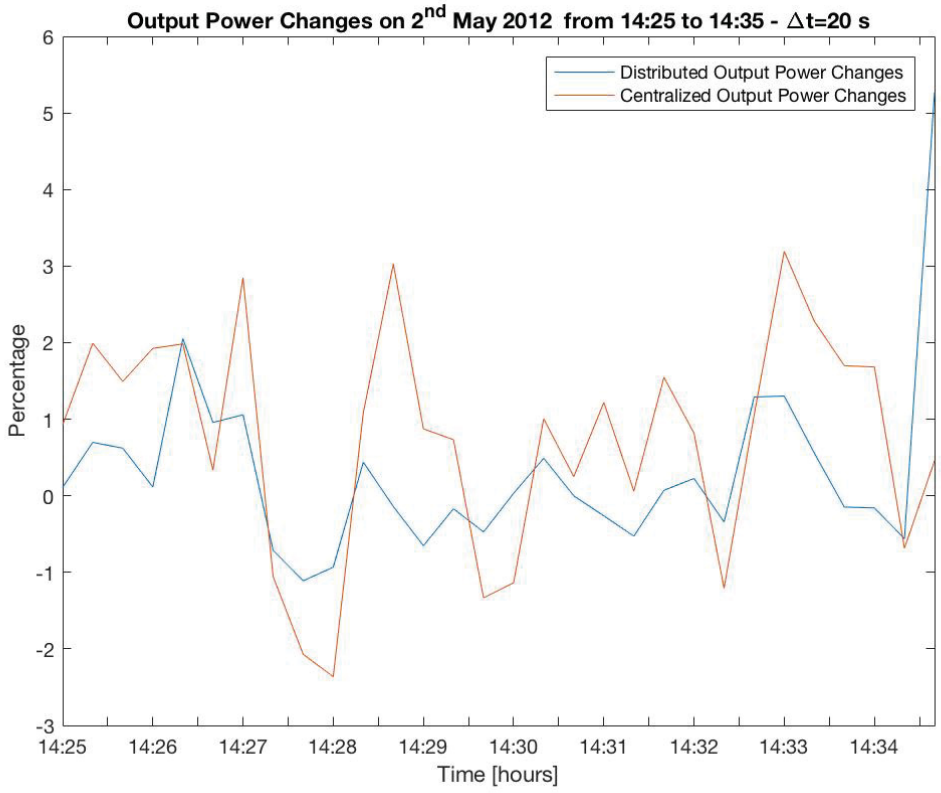

Fig. 23: $\Delta P_{\Delta t}$ evaluated over 10 minutes for both cases (centralized and distributed) considering a 5 MW power plant.

The main metric to quantify the Output Power Variability of a power plant is the standard deviation $\sigma_{\Delta t}$ of the normalized output power changes $\Delta \boldsymbol{P}$ (separated by the Time Interval Δt), as defined in (20).

$$\sigma_{\Delta t} = \sqrt{Var[\Delta \boldsymbol{P}]} \tag{20}$$

Therefore, the variability of a distributed power plant can be evaluated by considering a large number of scenarios. In all the performed simulations, the PV subsystems are

assumed identical (with the same capacity) and equally-spaced with a distance so that the output power time-series can be assumed uncorrelated.

The correlation between two adjacent subsystems can be calculated using the empirical model represented by equation (8) presented in [24] . Such relation connects the distance d between two locations (in meters) with the considered timescale Δt (in seconds) over which the variability is evaluated and with the average cloud speed v (in m/s). This correlation index decreases as the distance between locations increases and increases as the timescale of changes increases as demonstrated in [14]. It was also shown that a value lesser than 0.250 indicates a weak correlation between the time-series.

$$\rho = \frac{1}{1 + \frac{d}{v \cdot \Delta t}} \tag{21}$$

From now on the distance between two adjacent systems is set equal to 5 km. Such distance guarantees low correlation for all considered timescales in this thesis (from 2 to 120 seconds), relatively 0.005 and 0.238.

Therefore, the developed method has been used to calculate the standard deviation of the normalized output changes of different distributed power plants. In particular, the set of the analyzed capacities varies from 100 kW to 9.5 MW, power plants with such capacity range are typically studied in literature [17]. Moreover, the number N of subsystems ranges from 2 to 50. Let us remember that the distance between two adjacent subsystems is still considered equal to 5 km. It means that the extension of the distributed power plant can range from 5 and 245 km, PV power plant with such large sizes can be found in some studies as already described in the paragraph 2.2.

The plot of the standard deviation of the normalized output changes is shown in Fig. 24 by considering a Δt equal to 20 seconds and varying the total capacity of the power plant and the number N of subsystems over which the total capacity is equally divided, a higher N points out that more subsystems (but with smaller capacity) are employed in the power plant. The standard deviation is expressed in the percentage relatively to the power plant capacity in order to carry out a fair comparison, for example, the obtained value for considered power plant of 0.1 MW and distributed over two

locations has a standard deviation equal to 5.47%, it means that the actual standard deviation is 5.47 kW (5.47% of 0.1 MW).

It can be easily noted that, with equal power plant capacity, as the number of subsystems increases, the standard deviation decreases, this puts in evidence the smoothing caused by the geographical dispersion. Such effect is more evident when the power plant capacity is small, this can be explained by taking into account that, on bigger power plant there is already a stronger reduction of variability due to the physical extension of each location and therefore, the geographical dispersion has not a noticeable effect.

Moreover, with equal number of subsystems, the standard deviation decreases as the power plant capacity increases, such effect is less important when the number of subsystems is large because there is already a strong smoothing due to the geographical dispersion and in addition, the size of each subsystem is not enough large to provide a noticeable smoothing due to the size.

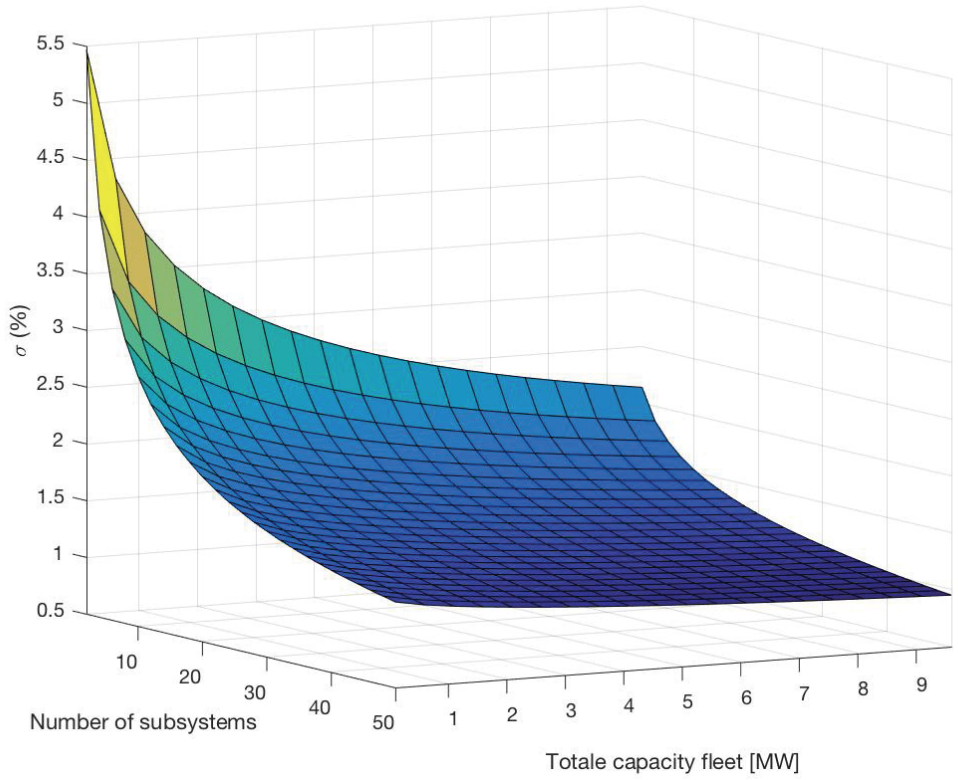

Fig. 24: Standard deviation of the normalized output power changes by varying power plant capacity and number of subsystems.

In other words, there are two kinds of smoothing that need to be considered when the evaluation of the standard deviation is performed, the one on each single subsystem due to its extension and another one due to geographical dispersion of the fleet, both effects have been considered by applying the low-pass filtering and by simply summing the output power resulting from the N subsystems.

3.5 Validation

In order to validate the results presented up to now, the same irradiance data were used to calculate the standard deviation of the output power changes of a distributed fleet, using the method proposed in [25].

The method uses the irradiance data, the number of subsystems N and the mean surface value $\bar{S}$ of subsystems in the fleet as inputs. In according with this method, it is possible to simulate the output power of the power plant applying two transfer functions in sequence. The first one simulates the low-pass filtering and produces the output power of a single location of the power plant. Such transfer function is the same that has been presented in [17]. The second one is useful to simulate the smoothing due to the geographical dispersion of the fleet over a region, it uses the number of subsystems N. Such transfer function, called $TF_{PV,fleet}$ is defined in (22) in the Laplace domain, where s is the Laplace variable, $p_N(s)$ is the output power of the fleet composed by N subsystems, $p_1(s)$ is the output power of a single location (already previously obtained by the first transfer function), and A is a coefficient which has been set equal to 2400 in the paper by fitting the real power data of the fleet located in Spain. Such method, with the same A value, has been applied and validated in the State of Colorado too, showing to work successfully in different climatic condition.

$$TF_{PV,fleet} = \frac{p_N(s)}{p_1(s)} = \frac{\left(A/\sqrt{N}\right)s + 1}{As + 1} \tag{22}$$

The Bode diagram for the proposed model is shown in Fig. 25. The red line represents the ratio between the real output power measured by two aggregated location and the real output power measured by a single location. The blue line does the same but considering the aggregation of five different locations. The proposed transfer function is represented by the superimposed black line, it puts in evidence the dependence on the number of the subsystems which composes the fleet, therefore as N increases, the fluctuations at high frequencies are reduced in magnitude.

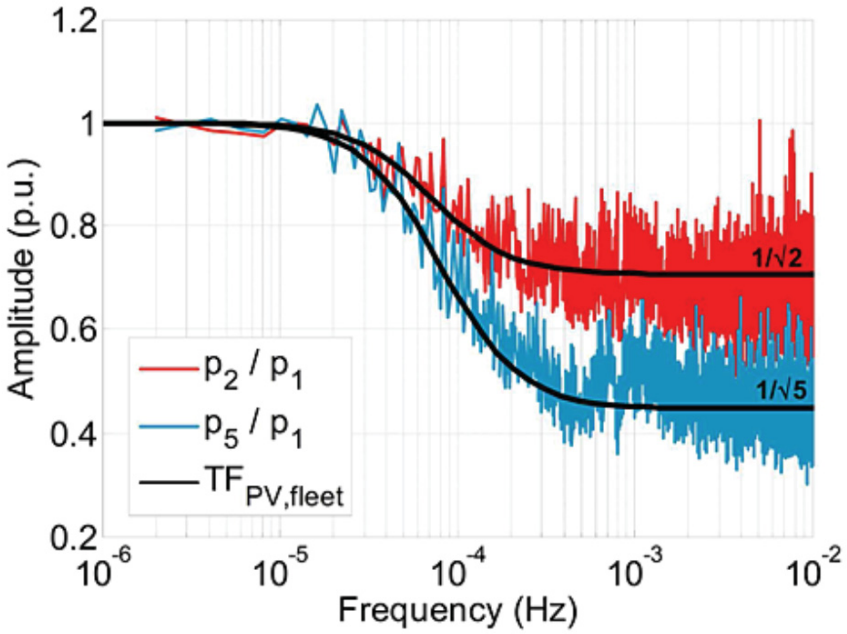

Fig. 25: Bode diagram of the model proposed in Eq.(22). At high frequencies, there is a reduction of variability which depends on the number N of aggregated subsystems [25].

At this point, the output variability has been quantified by calculating the differences between two consecutive output power samples, for a given Δt, as done previously and described by equation (19). Such operation has been done for both considered solutions (centralized and distributed).

The standard deviations of normalized output power fluctuations ΔP have been calculated (even now calculated with Δt equal to 20 seconds) and the resulting plot is shown in Fig. 26 by varying the power plant capacity and the number of the system of the distributed fleet. It has the same identical behavior of the results found previously, this puts in evidence the correctness of the obtained results.

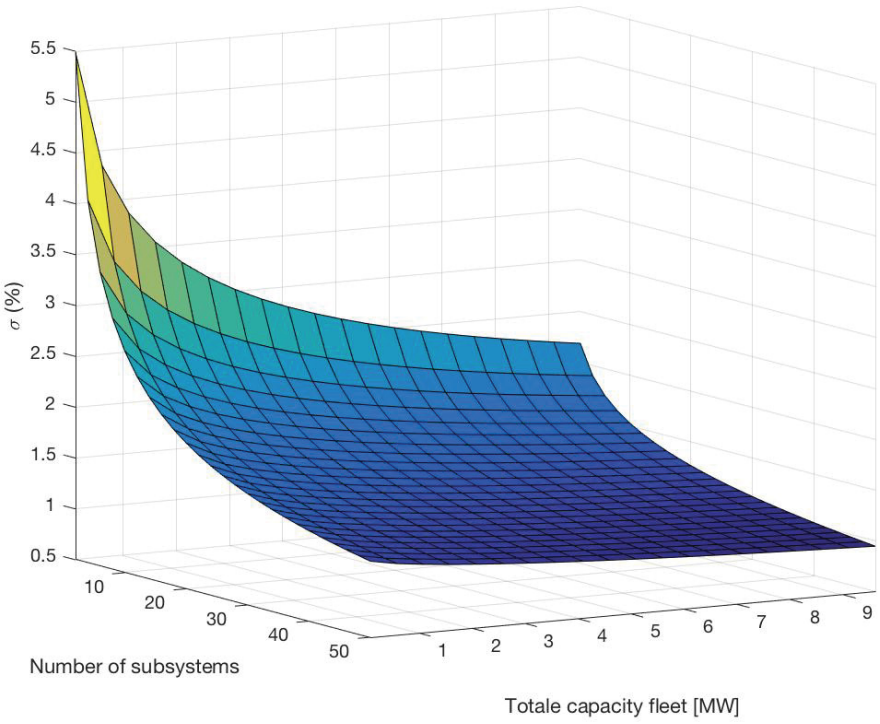

Fig. 26: Standard deviation of the Output Power Variability calculated by using the method described in [25].

To quantify the difference between the method proposed in [25] and the one developed in this thesis, the Root Mean Square Error (RMSE), defined in (23), has been calculated by considering the difference between the standard deviations varying C and N.

In (23), σ is the standard deviation calculated with the method presented in this thesis, while $\tilde{\sigma}$ is the same but calculated by using the method proposed in [25], C and N symbolize the capacity and the number of subsystems that vary, and n is the number of considered standard deviations.

$$RMSE = \sqrt{\frac{\sum_C \sum_N (\sigma - \tilde{\sigma})^2}{n}} \tag{23}$$

It has been found equal to 0.0115% confirming the goodness of the method proposed here.

Moreover, such error has been calculated for different Time Intervals (see Table 6), showing that the two methods agree on other short-timescales too, from 2 to 120 seconds, note that the order of magnitude of the error is very low (hundredth of a percentage point).

Table 6: Root Mean Square Error between the method proposed in [25] and the one presented in this thesis.

Time Interval [s]	2	5	10	20	30	60	120
RMSE [%]	0.0016	0.0035	0.0059	0.0115	0.0145	0.0413	0.1902

The accurate similitude between the results can be explained by considering that both methods take the smoothing effect on the single location (low-pass filtering) into account in the same way. Moreover, the second transfer function (22), which takes the geographical dispersion of the power plant into account, shows us that the smoothing (from the single location to the entire fleet) obtained at high frequencies is proportional to $\sqrt{N}$.

Such phenomenon is the same that can be observed in our case by evaluating the Relative Output Variability (ROV), already presented in the paragraph 2.3.3 and newly reported in (13), which is the ratio between the standard deviation of the output power changes of the entire fleet $\sigma_{\Delta t}^{Fleet}$ and the standard deviation of the output power changes of a single location $\sigma_{\Delta t}^{Subsystem}$.

$$ROV = \frac{\sigma_{\Delta t}^{Fleet}}{\sigma_{\Delta t}^{Subsystem}} \tag{24}$$

In fact, this quantity, evaluated by using the power data obtained with the method developed in this thesis, is always equal to $1/\sqrt{N}$, as shown in Fig. 27. This result is not surprising since that the locations can be considered uncorrelated with the considered Time Intervals. The Dispersion Factor (even it presented in the paragraph 2.3.3), considering that the systems are 5 km far away among them, a Time Interval equal to 20 seconds and that the cloud speed is 13 m/s, is calculated in (25) and it can be seen that it is much greater than N (ranging from two onwards), therefore D is in Spacious Region.

$$D = \frac{L}{v \cdot \Delta t} = \frac{5000 \cdot (N-1)}{13 \cdot 20} = 19.2(N-1) \gg N \tag{25}$$

This matches with the results presented previously in 2.3.3.

Let us remember that the Relative Output Variability is related with the smoothing due to the geographical dispersion of the fleet, in particular, such quantity and the smoothing are inversely proportional, as ROV decreases, the reduction of variability from a single subsystem to the entire fleet is more remarkable. Therefore, this relation, verified in our case, agrees with the Spanish model because ROV is equal to $1/\sqrt{N}$, whereas in (22) the smoothing is proportional to $\sqrt{N}$.

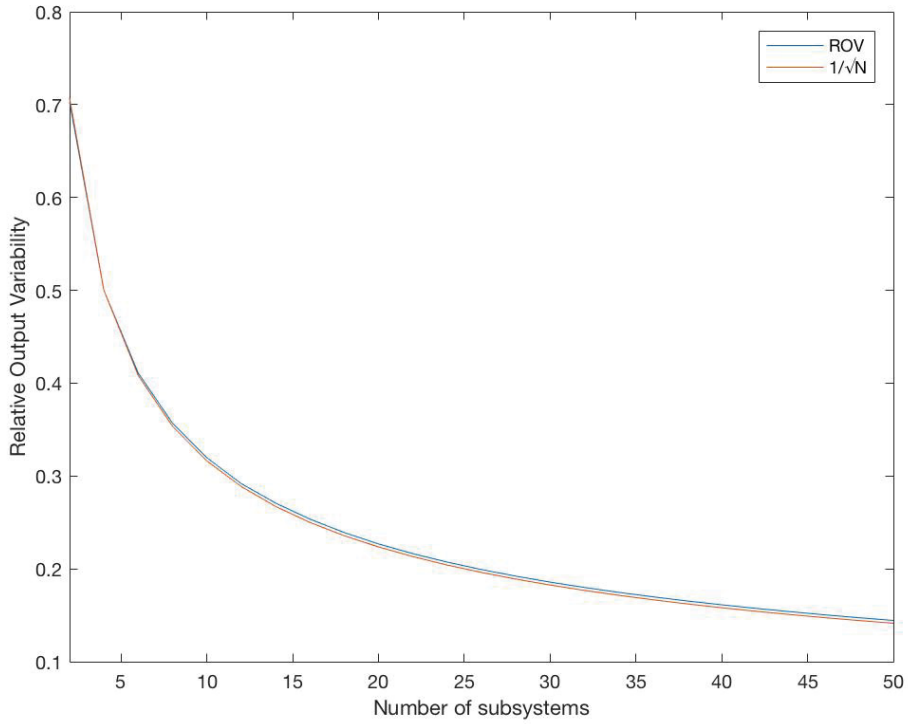

Fig. 27: Relative Output Variability evaluated by using the simulated output power data.

In the method developed in this thesis (see Fig. 20), by shifting the irradiance data and summing the output power of the different virtual locations, the output power of the distributed fleet is obtained. The results point out that, by summing the different outputs, there is a remarkable compensation of the fluctuations in the output of the different locations. Such phenomenon is the same that happens in the frequency domain as shown in the Bode diagram in Fig. 25. Therefore, probably if the measurements were done on a real power plant, a similar result would be obtained.

Note that, the curve represented in Fig. 27 has been calculated by considering the output power, not the irradiance data as done in the paragraph 2.3.3. This agreement

between the results points out that the smoothing effect on a single location has a very negligible effect on the Relative Output Variability evaluation. More clearly, the ROV has quite exactly the same value in both considered cases, by calculating the standard deviation of the output power changes or the standard deviation of the irradiance data changes. Therefore, it has been proved that the smoothing effect on each subsystem of the fleet can be left out in the evaluation of the ROV.

3.6 Final consideration

In conclusion, in this chapter, it has been shown how the evaluation of the output power variability has been performed in this thesis. In particular, by using the method to calculate the simulated output power of a single location (just knowing the irradiance point measurement and the system size), a method to estimate the output variability of a generic power plant has been developed. Successively, the results have been matched with some scientific works present in literature, that allows us to validate the developed method.

The two-main modelled effects are the local smoothing on each subsystem due to its physical size and the global smoothing due to the extension of the fleet over a large geographical region.

In addition, a substantial equivalence has been found in the ROV calculation, when the irradiance data are used instead of the power data.

In the next chapter, a comparison between two kinds of power plant will be performed, analysing how the two smoothing effects influence the power fluctuations by varying the system capacity.

4 A comparison among different footprints

In this chapter, the main aim is to perform an analysis of the Output Power Variability (estimated in the chapter 3) and make a comparison between two kinds of footprints for a PV power plant (centralized and distributed) in order to understand how the variability is affected by the capacity of the power plant in the two different layouts.
Different metrics have been employed, like standard deviation and largest fluctuation with the goal of understanding which solution provides the more stable output power and quantifying this difference in relation with the power plant capacity, the number of subsystems over which the power plant is distributed and the considered Time Interval.
To better understand, the comparison is done by analyzing the Output Power of two kinds of layouts, as shown in Fig. 28. Such figure displays two PV power plants with the same total capacity (in this case 5 MW) but organized in different ways. Fig. 28a) shows a power plant composed by a single PV system located in a single location, while Fig. 28b) shows five different subsystems, each one has a capacity equal to 1 MW and they are supposed to be located along the main direction of movement of shadows caused by moving clouds, as already mentioned in the paragraph 3.1. The differences between the two solutions are studied, at the beginning, by varying the power plant capacity to figure out how the local smoothing due to the physical extension of each subsystem has an effect on the variability, and successfully the number of subsystems of the distributed solution varies too, that is done to understand how the geographical dispersion influences the smoothing effect.

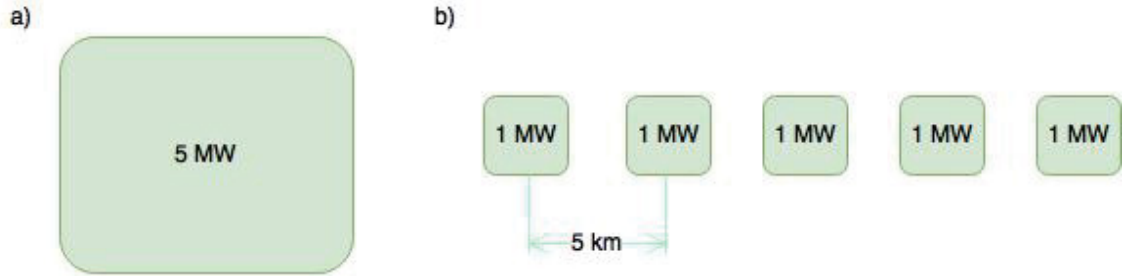

Fig. 28: a) A centralized power plant of 5 MW and b) a distributed PV fleet, in which the same capacity (5 MW) is divided on 5 subsystems, each one of 1 MW.

4.1 Centralized vs. Distributed

In a first simulation, the set of the analyzed capacities varies from 100 kW to 9.5 MW, while for the distributed fleet, the number N of subsystems was set equal to 5.

The comparison has been done by setting the capacity of each subsystem equal to the same of the centralized one, divided by N, so the comparison is done by considering the same total capacity. The Δt over which the variability of output changes is evaluated was set equal to 20 seconds.

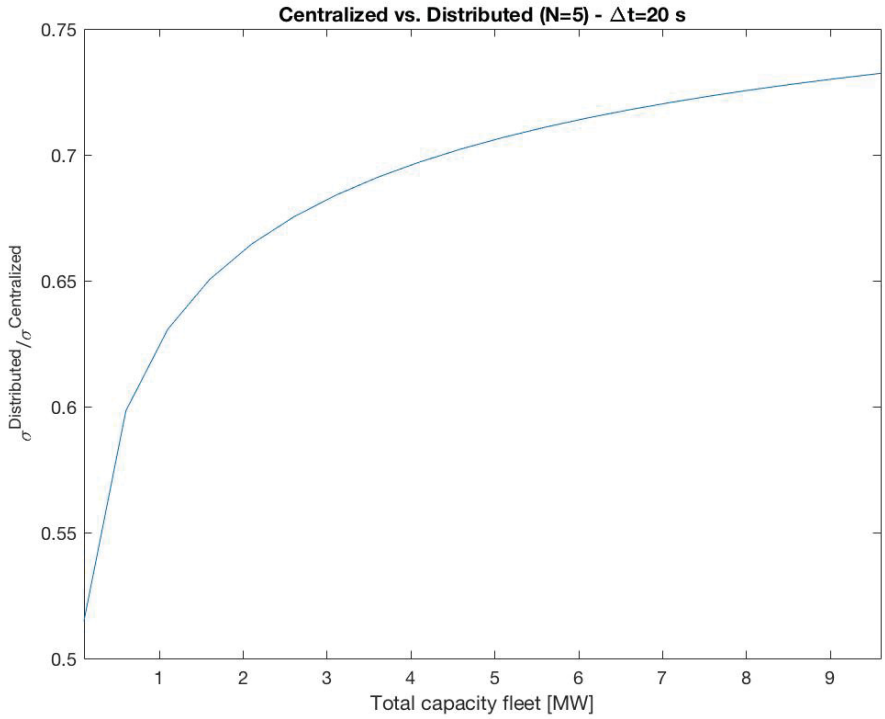

Fig. 29: Reduction of standard deviation of output changes from a centralized to a distributed system, by varying the capacity of the fleet.

Fig. 29 shows the ratio between the standard deviations of the normalized output changes of the distributed and centralized solutions. For example, it can be seen that for a 5 MW PV power plant, this ratio is equal to 0.7, this means that the standard deviation of the distributed solution (denoted $\sigma^{Distributed}$) is the 70% of the standard deviation of the centralized one (denoted $\sigma^{Centralized}$), therefore the variability of the distributed power plant is minor than the variability of the centralized power plant.

By using the Matlab® Curve Fitting Toolbox, a fitting operation was done in order to find a relation between the reduction of variability (represented by the ratio of the standard deviations) and the power plant capacity C. The found relation is reported in

(26), it is a power function with positive exponent, where C identifies the power plant capacity expressed in MW, in this specific case (N equal to 5) a is equal to 0.62 and b is 0.07.

$$\frac{\sigma^{Distributed}}{\sigma^{Centralized}} = aC^b \quad (26)$$

It can be highlighted that, as the dimension of the fleet increases, the ratio increases too, so the advantages of a distributed solution are, even though still relevant, reduced.

This can be explained considering that, when the size of a centralized solution is smaller, the local smoothing effect is quite weak (because of higher cut-off frequency) so the geographic dispersion has a stronger impact on the global evaluation of the standard deviation of the power plant. As the power plant size increases, the mitigation of fluctuations, already in the centralized solution, becomes stronger because of the physical extension of the plant and so the geographic dispersion cannot improve the reduction so much and therefore there is a limited variability reduction.

The largest power fluctuation ΔP (considering the absolute value) has been calculated for both cases and for different capacities and the results are shown in Fig. 30. There is a great difference between the values in the case of a 100 kW plant, showing an important impact of the geographical dispersion on this quantity considering the smaller systems. As the system capacity increases this difference becomes quite negligible, and we can even see that, for 9 MW analysis, the distributed solution produces a larger fluctuation than the centralized one. A possible reason is that the lower cut-off frequency (caused by only one larger system) affects this quantity more than the combination of the local smoothing (on each smaller PV system of the plant) and the geographical smoothing.

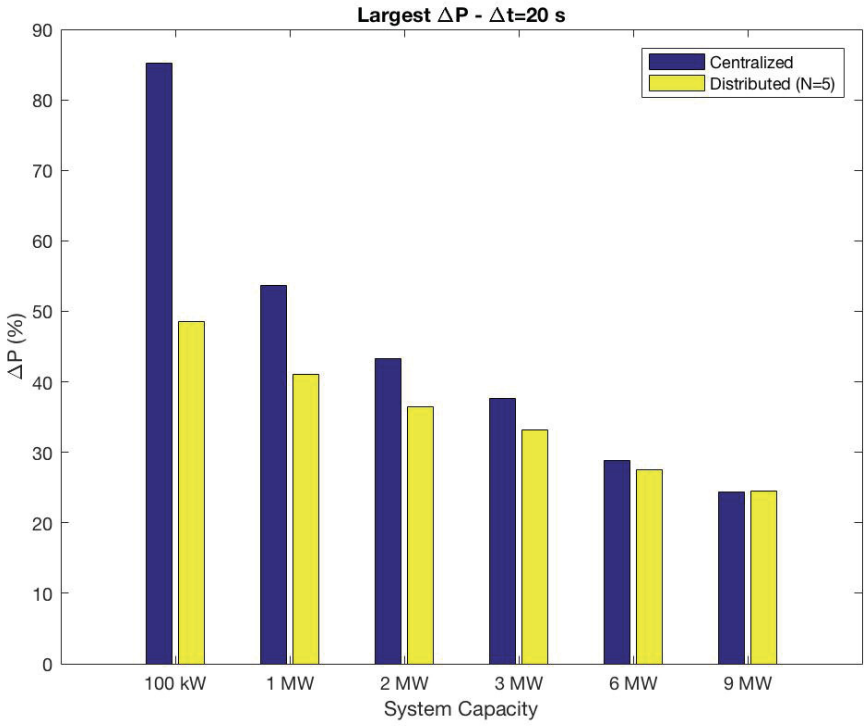

Fig. 30: Comparison about the largest ΔP in percentage for different capacities

Another metric that it is of interest for the grid planner is the number of ΔP values that are higher than 10% of the power plant capacity. In other words, this quantity means the percentage of total fluctuations which goes beyond 10% of the capacity during a single Time Interval. In Fig. 31 this amount is shown in percentage, it can be observed that, in the case of a distributed plant, the number of fluctuations is always lower than for the case of a centralized solution. This points out that the probability to have a large output fluctuation is lower if a distributed solution is adopted.

Note that, as the system capacity increases, the difference between the two solutions becomes more evident. For a 1 MW centralized system, 2.1% of the fluctuations exceeds 0.1 MW (that is 10% of 1 MW) during the considered Time Interval (20 seconds), while for a 1 MW distributed plant, just 1% of the ΔP values exceeds 0.1 MW. While considering a system with a capacity of 9 MW, the improvement, by spreading the power plant over 5 subsystems, is from 0.9% to 0.2%. Therefore, in according with this quantity, and in contrast with the analysis done on the standard deviation previously, the improvement is more appreciable as the power plant capacity is larger.

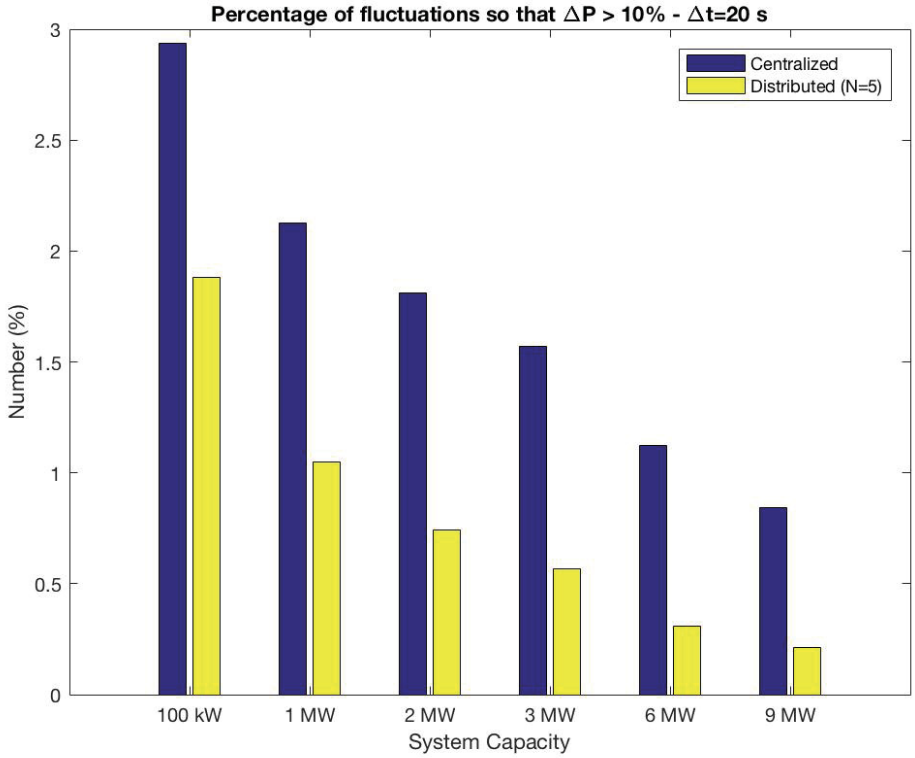

Fig. 31: Comparison between the percentage of fluctuations so that $\Delta P > 10\%$.

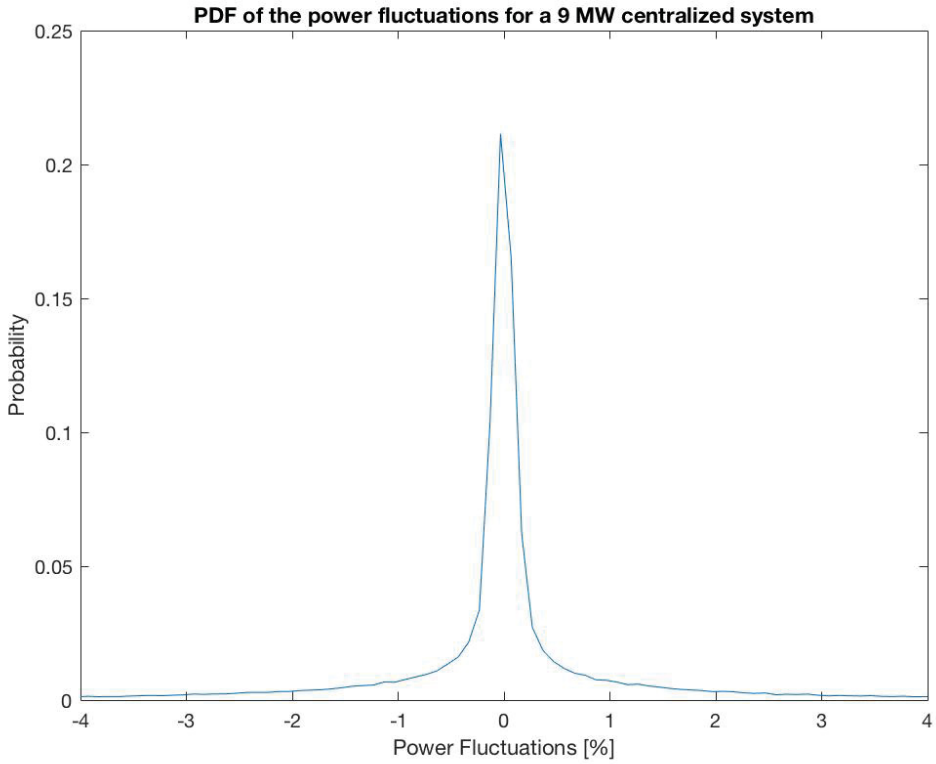

Fig. 32: Probability Density Function (PDF) of the power fluctuations in the case of a 9 MW centralized system.

Fig. 32 shows the Probability Density Function (PDF) of ΔP distribution for a 9 MW centralized plant, such figure puts in evidence that the random variable ΔP does not follow a Gaussian distribution since that it exhibits a pronounced peak and long tails. The shape of the distribution shows a higher probability of extreme fluctuations than would be expected for a Gaussian distribution with a similar standard deviation. In other words, the distribution of the power fluctuations exhibits "fat tails" relative to a Gaussian distribution. To describe its similitude to the normal distribution, it is used a metric called $k_{3\sigma}$ [14], defined in (27), that is the ratio between the 99.7th percentile of the output changes and the standard deviation of the distribution. A high value of $k_{3\sigma}$ points to a higher probability of having a large output changes.

$$k_{3\sigma} = \frac{99.7^{th}(\Delta P)}{\sigma_{\Delta P}} \tag{27}$$

For a normally distributed variable the $k_{3\sigma}$ value should be almost equal to 3. This would mean that three times standard deviations contain the 99.7% of the values.

Fig. 33 shows the $k_{3\sigma}$ values for different power plant capacities. For centralized plants, $k_{3\sigma}$ decreases with increasing system size, especially from 100 kW to 1 MW. It means that the number of extreme fluctuations decreases with increasing PV plant size due to higher cut-off frequency of the low-pass filtering that smooths the more extreme fluctuations. In addition, by increasing PV plant size, the PDF is always more similar to a Gaussian since that this parameter decreases.

Whereas, for distributed solution, $k_{3\sigma}$ is always equal to 6, not showing an important improvement by changing capacities, so the power plant capacity does not seem influence the relation between the 99.7th percentile and the standard deviation. That means that the 99.7th percentile of the fluctuations distribution does not exceed six times standard deviation. Since that this value keeps constant as the capacity increases, the probability of having fluctuations which exceed six times standard deviation (let us remember that the standard deviation is a quantity already normalized to the plant capacity) remains the same.

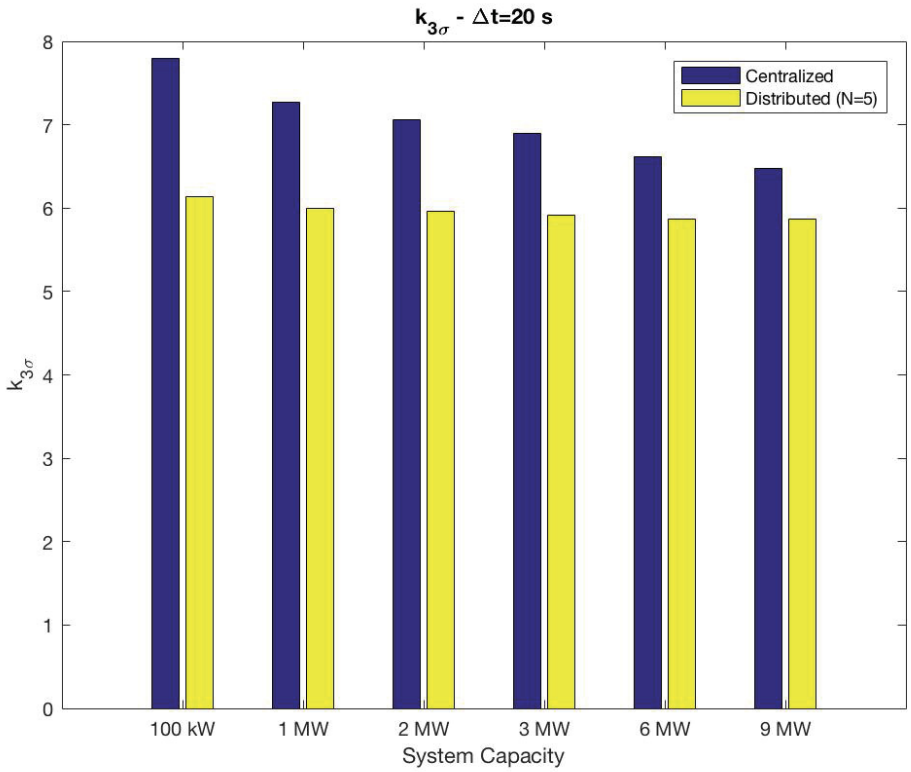

Fig. 33: $k_{3\sigma}$ value for both solutions (centralized and distributed), considering different power plant capacities.

4.2 Effects of changing the number of subsystems

It is time to consider what happens by changing the number of subsystems over which the capacity of the fleet is split in the distributed footprint. Over the following analysis, the number of subsystems ranges from 2 to 50.

Fig. 34 and Fig. 35 show the Probability Density Function (PDF) respectively considering 0.1 MW and 9 MW as system capacity and setting Δt equal to 20 seconds. It can be clearly seen that the standard deviation decrease as the number of subsystems N increases, this highlights the effects of the geographical smoothing.

In addition, it can be observed that, by observing the X-axis, the magnitude of fluctuations in the case of 0.1 MW plant is greater than the 9 MW plant (by considering equal N) and this is true for both solutions (centralized and distributed), showing an important reduction of variability due to the size of the considered systems.

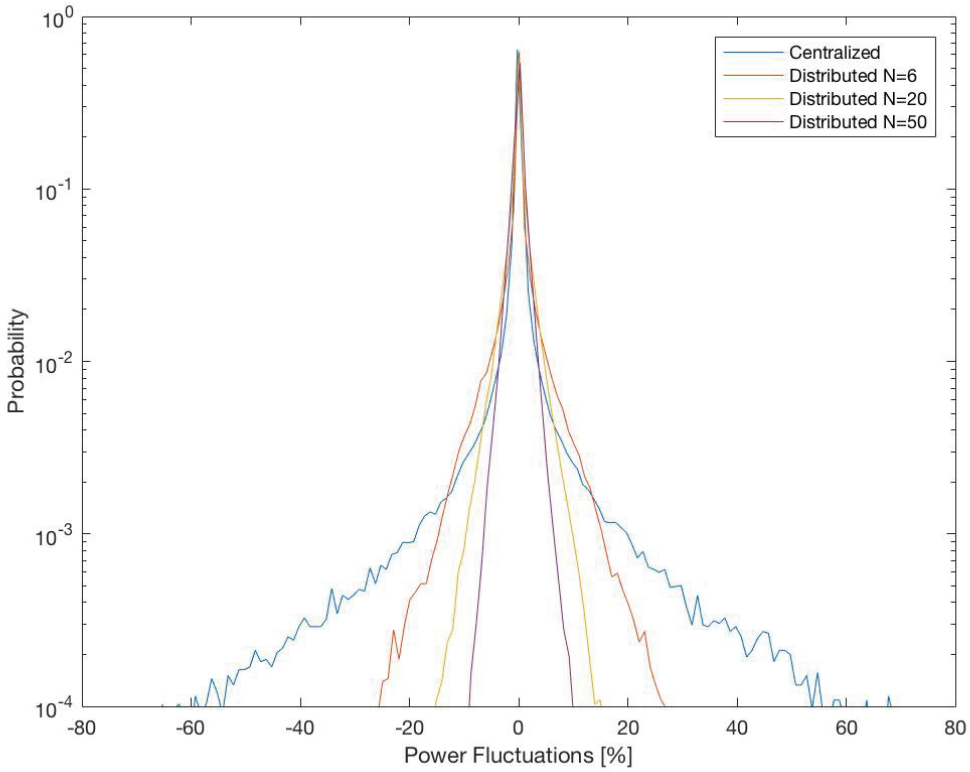

Fig. 34: Distributions of power fluctuations on a 100 kW PV system for different N.

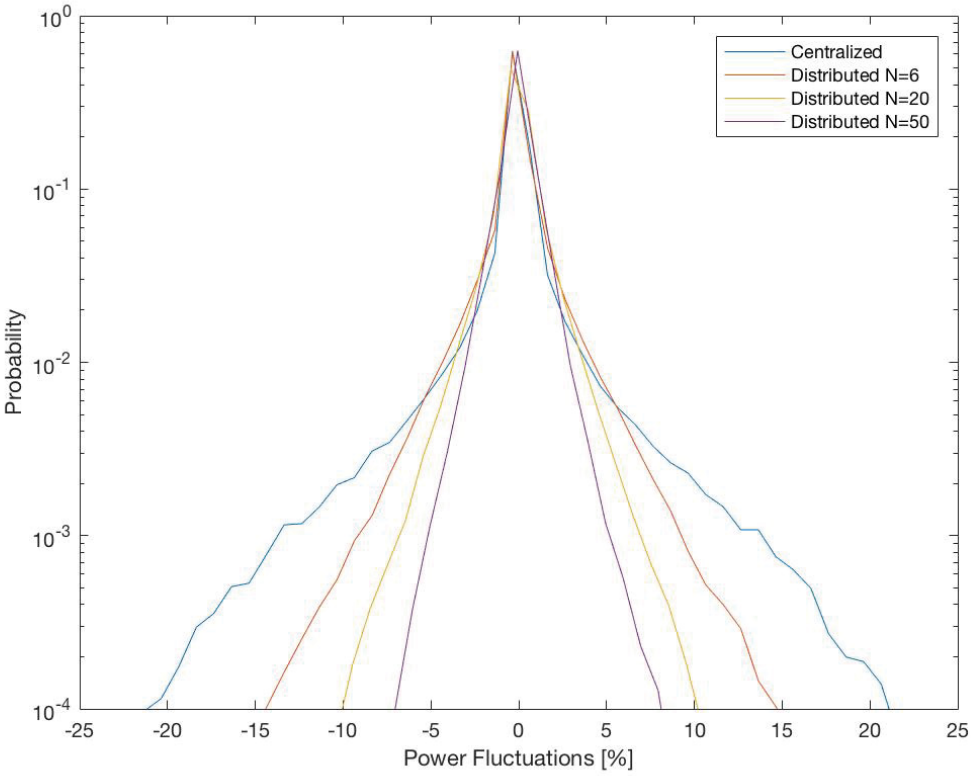

Fig. 35: Distributions of power fluctuations for a 9 MW system considering different N.

Table 7: Percentage of ΔP intervals considering the capacity equal to 0.1 and 9 MW and different number of subsystem.

ΔP(%)	Capacity [MW]							
	0.1				9			
	N=1	N=6	N=20	N=50	N=1	N=6	N=20	N=50
$0 \leq \Delta P < 1$	88.67	85.62	85.14	87.43	89.85	88.50	89.07	90.12
$1 \leq \Delta P < 3$	4.78	6.78	10.04	10.62	4.66	7.60	8.73	9.03
$3 \leq \Delta P < 10$	3.58	6.03	4.62	1.94	4.52	3.72	2.19	0.84
$10 \leq \Delta P < 20$	1.51	1.37	0.19	0.01	0.91	0.17	0.01	0.01
$20 \leq \Delta P < 30$	0.67	0.17	0.01	0.00	0.05	0.01	0.00	0.00
$30 \leq \Delta P < 100$	0.79	0.03	0.00	0.00	0.01	0.00	0.00	0.00

Table 7 shows how some fluctuation intervals are distributed considering the two borderline cases of this study, that are 0.1 and 9 MW. This table analyses the influence of PV system capacity and number of subsystems on the PDF of the fluctuations, putting in evidence that as the capacity and number of subsystems increase, the probability of having more extreme fluctuations decreases.

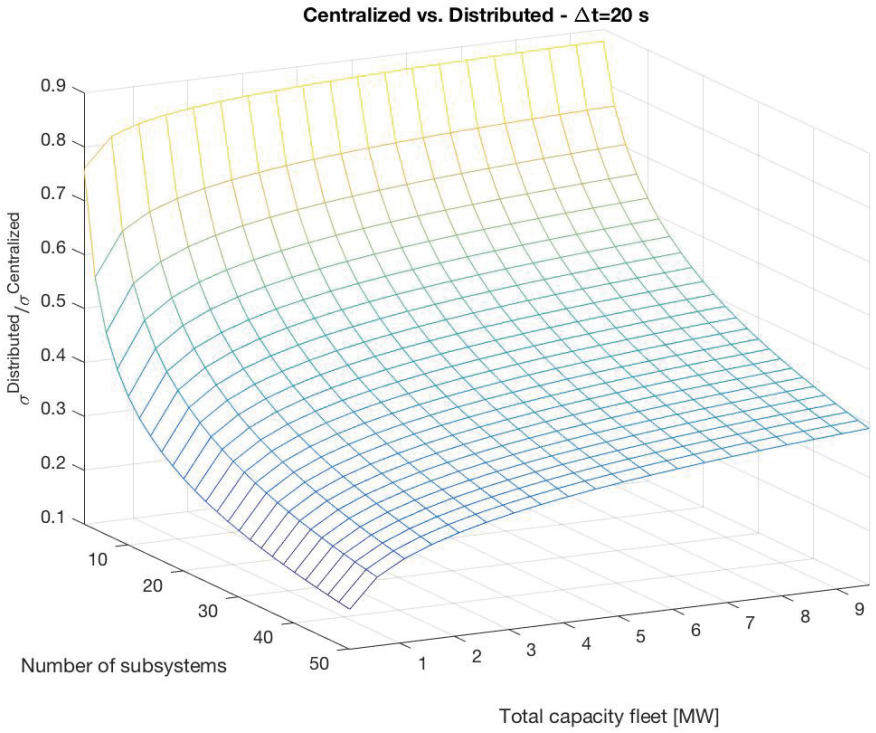

Fig. 36: Ratio of the standard deviation of the output power changes of a distributed PV system with respect to a centralized system as a function of the number of subsystems and the total power production capacity. The Time Interval Δt was 20 s.

Fig. 36 shows the ratio between the standard deviations of ΔP of the distributed and centralized solutions by varying the capacity of the fleet and the number of subsystems that compose the fleet in the distributed solution. Such ratio quantifies the reduction of variability passing from a centralized to a distributed solution.

It can be observed that by increasing the number of subsystems, there is an important reduction of the magnitude of the fluctuations (so an improvement) because the standard deviation of the entire distributed plant decreases, this is due to the geographical distribution of the fleet.

For a fixed N, obviously, the mathematical relation is the same reported in Eq. (26). The coefficient a and b are reported in Table 8 for some cases and it can be noted that, as N increases, a decreases and b increases.

The decreasing rate of the ratio is higher when N is very low, this kind of behavior is shown for all the considered capacities. For example, by subdividing 100 kW from 2 to 6 subsystems, there is an improvement of 38% on reduction of variability (shown in Fig. 36), while from 2 to 50 subsystems the improvement is equal to 77%. Whereas, considering a total capacity of 9.5 MW and subdividing this capacity from 2 to 6 subsystems, the improvement is just of 20%, much smaller than 100 kW plant case, and from 2 to 50 there is an improvement of 55%.

Table 9 summarizes the results, for Δt equal to 20 seconds, obtained considering different capacities and different number of subsystems.

It points out that, fractionating the power plant into more subsystems takes greater advantages for small power plants in respect to greater ones. This can be explained by considering that, when the plant size is smaller, the smoothing effect due to the size of the system is practically absent, so the geographical dispersion can give a more important contribute on the reduction of variability.

Table 8: Coefficient a and b for the equation (26) by varying N.

Number of subsystems	a	b
2	0.82	0.03
6	0.58	0.08
10	0.49	0.11
30	0.31	0.16
50	0.25	0.19

Table 9: Reduction of standard deviation (in percentage) by increasing the number of subsystems totalizing different capacities.

Number of subsystems	Capacity [MW]				
	0.1	1	3	6	9.5
From 2 to 6	38%	29%	24%	21%	20%
From 6 to 50	64%	57%	52%	47%	45%
From 2 to 50	77%	69%	63%	58%	55%

Fig. 37 shows the largest power fluctuation ΔP for different capacities and for different configurations (i.e. different numbers of subsystems).

The general trend is that, as the number of subsystems N increases, the maximum ΔP decreases. Notice that, in accordance with this quantity, concentrating all the capacity in a single location can be a better solution than fractionating that into two locations if we consider a power plant equal or greater than 1 MW. This can be explained considering that the smoothing effect due to the size of a large plant is stronger than the sum of effects due to the size of two plants (totalizing equivalent capacity) and geographical dispersion.

Moreover, we can observe that varying the capacity of the plant gives just a negligible reduction on the maximum ΔP in the case of the fleet is distributed over many subsystems (with N equal to 20 or 50), this points out that the smoothing effect due to the subsystem size does not influence the largest fluctuation so much when N is large. This can be explained by considering that in that case the size of the physical size is too small, and therefore the cut-off frequency is quite large.

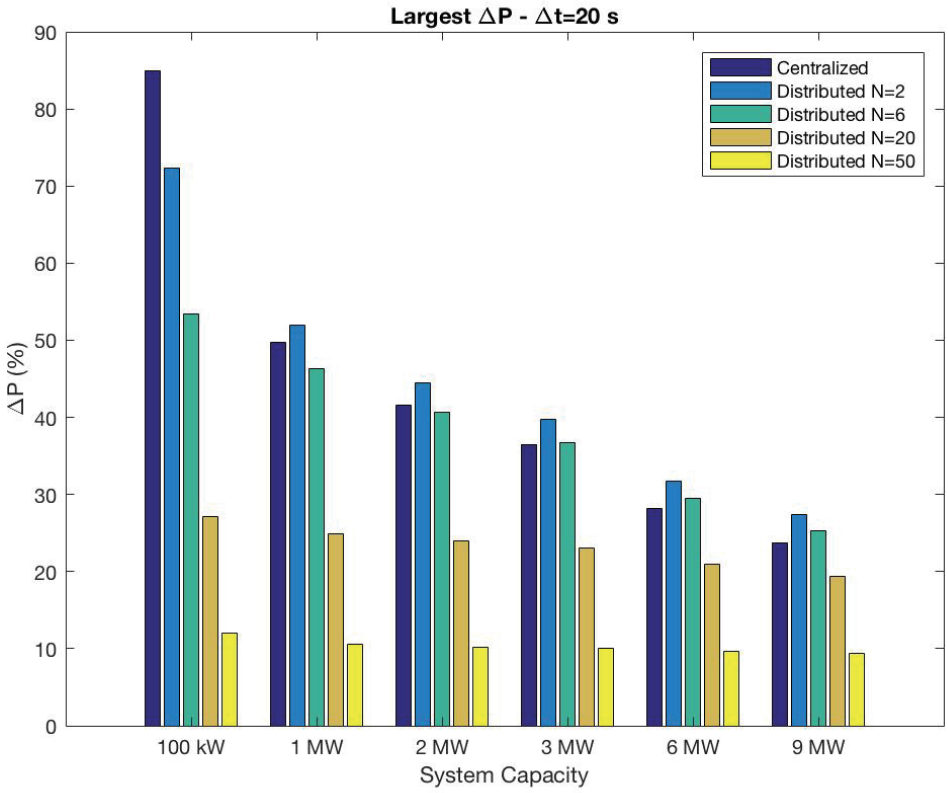

Fig. 37: Comparison about the largest relative ΔP by varying capacity and number of subsystems.

Fig. 38 shows the percentage of fluctuations that are larger than 10% of the capacity, as already done previously simply by considering the 5 MW case. It can be seen that there is always an improvement by increasing the number of subsystems, in particular the percentage with 6 subsystems is, more or less, the half of the one with 2 subsystems, showing an important improvement. In addition, with N is equal to 50, the percentage of fluctuations larger than 10% are practically absent, showing that geographical smoothing (over many subsystems) has a very strong effect on the reduction of fluctuations.

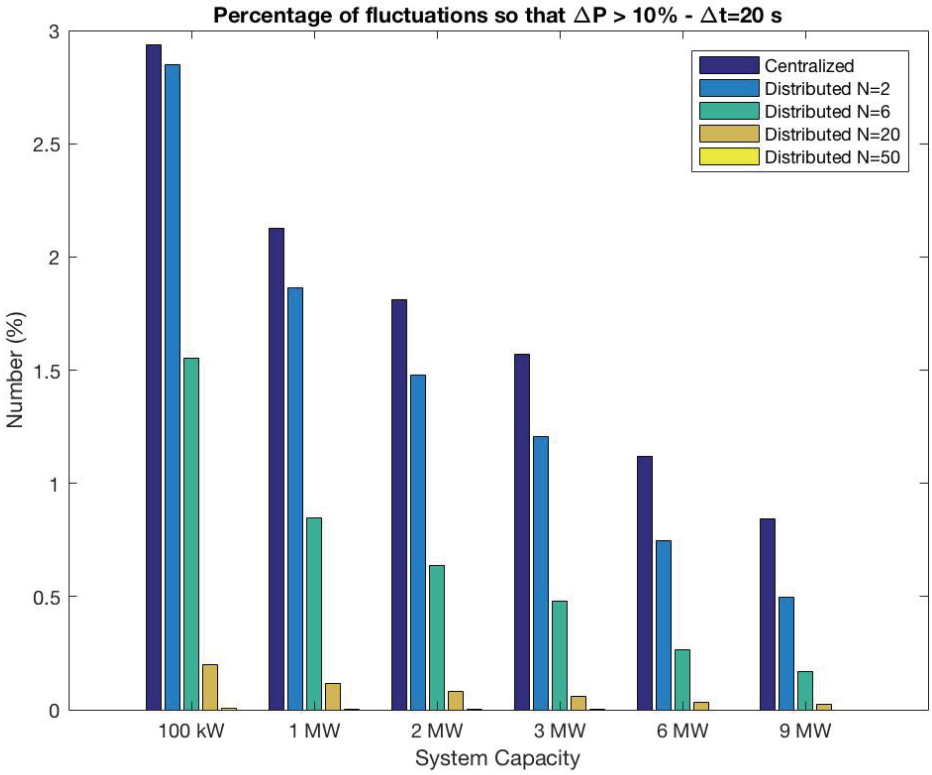

Fig. 38: Share of PV power fluctuations with ΔP > 10%, by varying capacity and number of subsystems.

Fig. 39 shows the $k_{3\sigma}$ values, defined in (27), for different capacities and for different number of subsystems. For the centralized solution, there is an improving, especially from 100 kW to 1 MW, of this parameter. This is due to the cut-off frequency (given by a $1/\sqrt{S}$ law) of the low-pass filtering that smooths the more extreme fluctuations.

Whereas, for distributed solutions, $k_{3\sigma}$ is quite always the same value for every total capacity (with equal N), not showing any important decrease by changing capacities. Therefore, the power plant capacity does not seem influence the relation between the 99.7th percentile and the standard deviation of the PDF.

This indicates that the probability of having a fluctuation which exceeds the 99.7th percentile of the distribution is the same by varying the power plant capacity.

By changing the number N of subsystems, with equal total capacity, this quantity decreases putting in evidence the geographical dispersion effect. The physical meaning is that the 99.7th percentile of the distribution is closer to the standard deviation, so the probability of having an extreme power fluctuation is lower.

In conclusion, the smoothing on each subsystem has a negligible effect on the evaluation of this quantity, while the geographical dispersion has a noticeable effect, because as N increases, $k_{3\sigma}$ decreases.

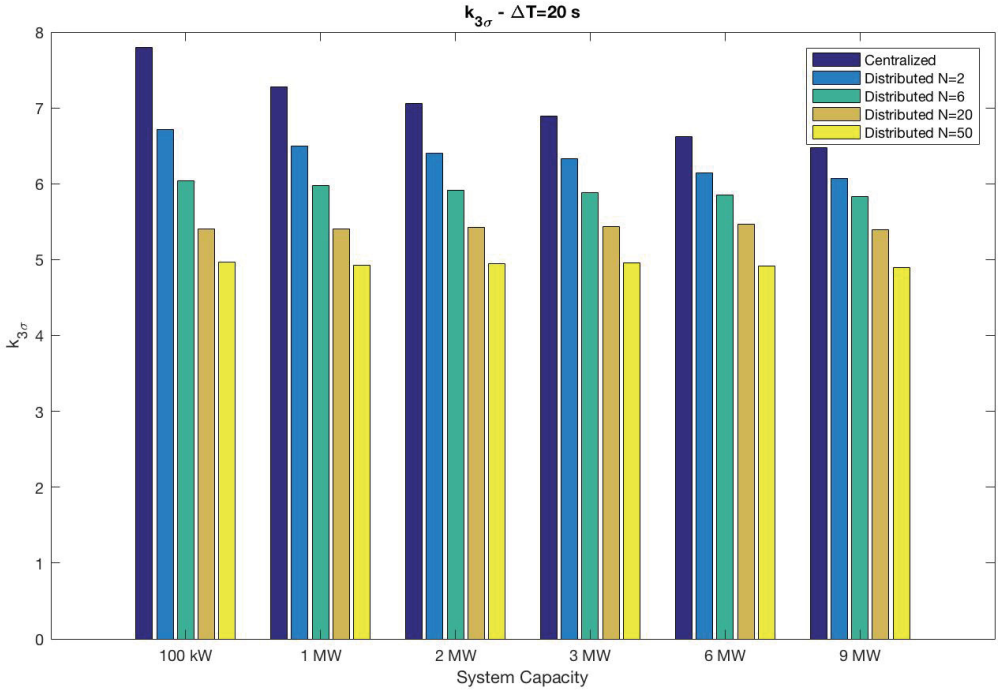

Fig. 39: $k_{3\sigma}$ value considering different capacities and different number of subsystems.

4.3 Effects of changing the Time Interval Δt

Now, the influence of Time Interval Δt can be considered on the output power variability as the capacity of the power plant varies. Let us remember that Δt specifies the time interval between two consecutive output power values used to calculate the fluctuations ΔP. During the following simulations, concerning the distributed solution, the number of subsystems, over which the total capacity is distributed, is set equal to 5. The analyzed Time Intervals ranges from 2 seconds to 120 seconds, such timescales are typical for urban PV applications.

Fig. 40 and Fig. 41 show the Probability Density Function of power fluctuations of centralized and distributed plants, respectively sizing 100 kW and 9.5 MW analyzed considering different timescales. It can be seen from the plots, as the Time Interval increases, the probability of having stronger power fluctuations increases too, these

results agree with the ones shown in the paragraph 2.1 and discussed in detail in [5]. In addition, it can be noted that the standard deviation decreases as the dispersion of the power plant increases (from one single location to five locations) and this is true for all considered Time Interval.

Moreover, by comparing Fig. 40 and Fig. 41, the standard deviation decreases as the power plant capacity increases as previously underlined in chapter 3, this is true especially for lower timescales, while for long timescales (like Δt = 120 s), this reduction is very limited.

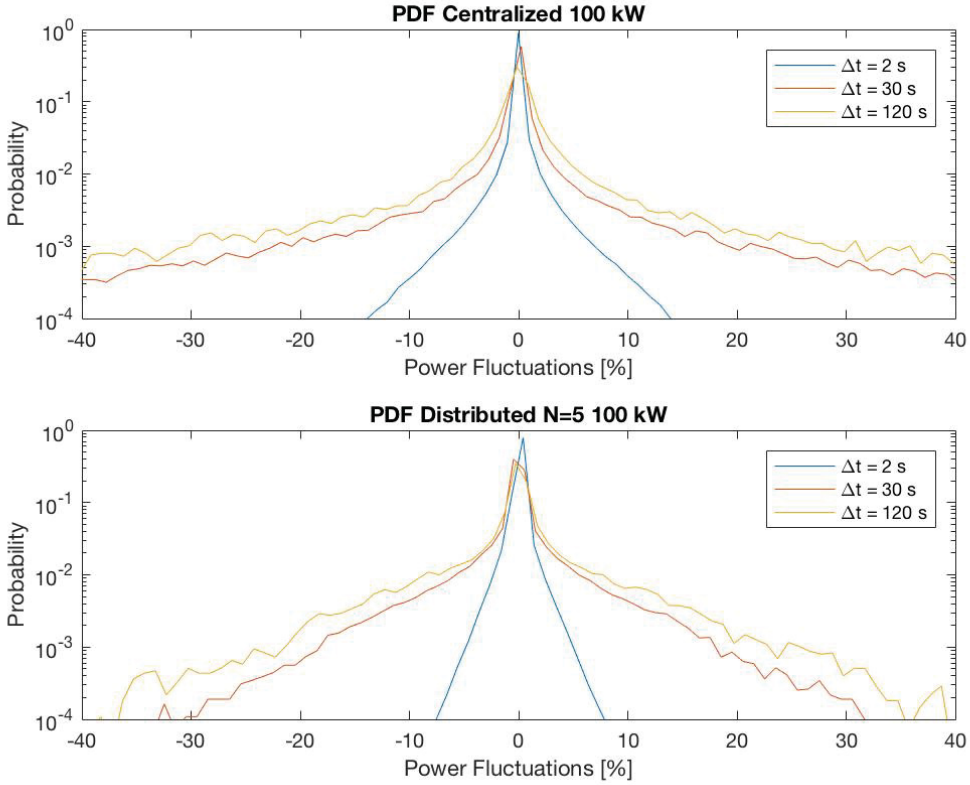

Fig. 40: Distributions of power fluctuations on a 100 kW power plant considering different time scales and both scenarios (centralized and distributed).

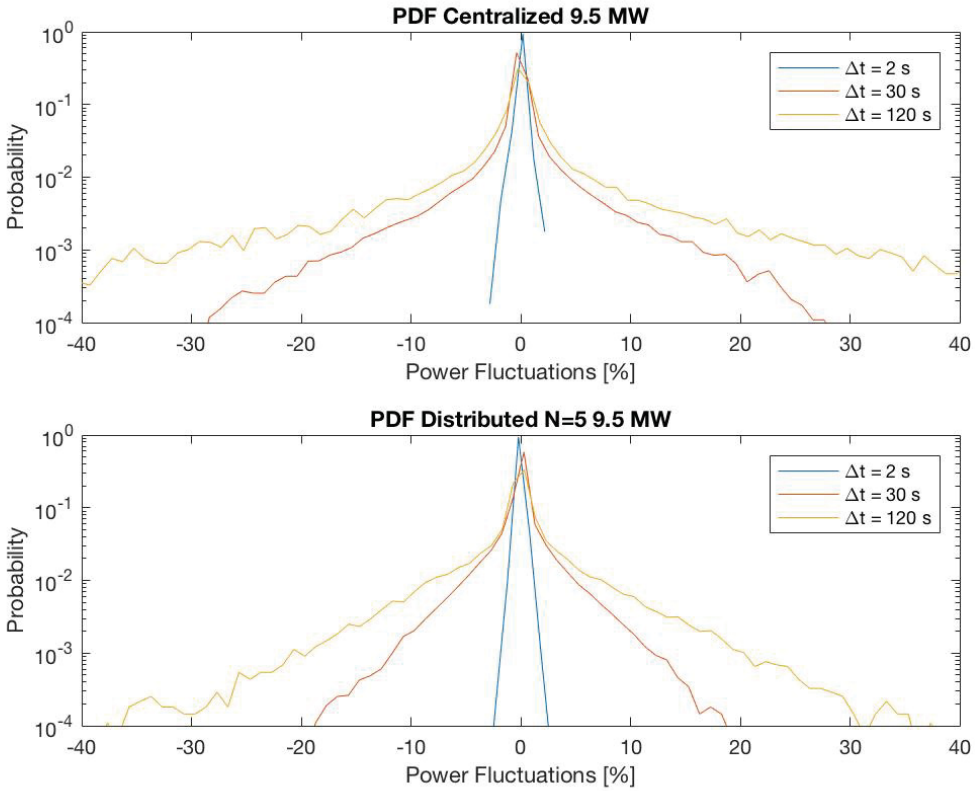

Fig. 41: Distributions of power fluctuations for a 9.5 MW system considering different time scales and both scenarios (centralized and distributed)

Fig. 42 shows the reduction of variability, quantified in according with the standard deviation, switching from a centralized power plant to a distributed one (five uncorrelated subsystems). It can be observed that this ratio decreases as the Time Interval increases, therefore, the geographical dispersion of a PV power plant takes more advantages with longer timescales. Such trend is valid for all the considered capacities.

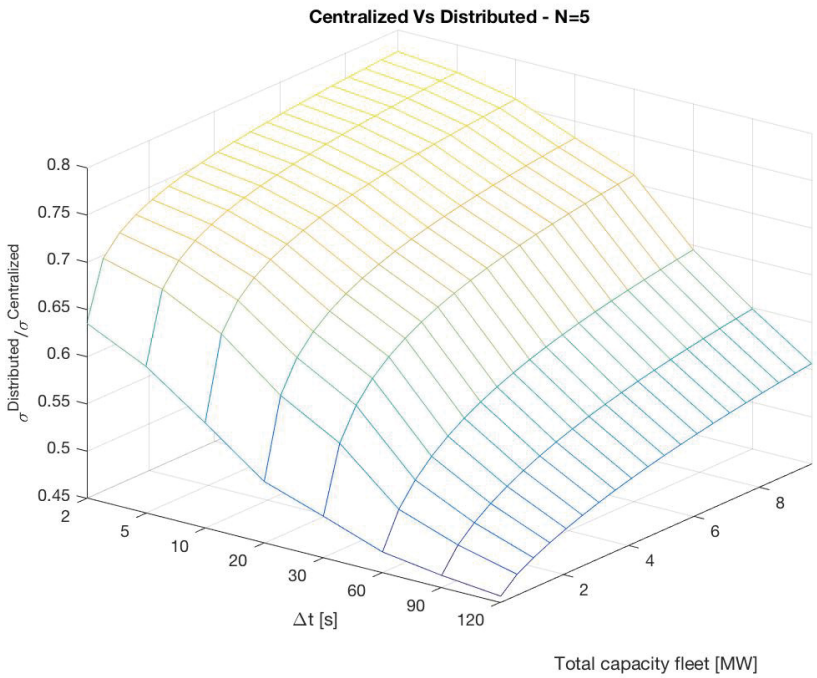

Fig. 42: Reduction of standard deviation of output changes from a centralized to a distributed fleet (N=5), by varying the capacity of the fleet and the Time Interval Δt.

Table 10 displays the influence of changing the timescale on the reduction of variability considering different capacities. It puts in evidence that in the case of a 0.1 MW the more important improvements happen at shorter timescales (from 2 to 30 seconds), with minor improvements considering larger timescales. The decreasing rate of the ratio has a similar behavior for the other capacities, but with a negligible difference at shorter timescales, as we can see in the last columns of the table, and a strong improvement with larger Time Interval.

Table 10: Percentage of improvement on $\sigma^{Distributed}/\sigma^{Centralized}$ by increasing the time interval considering different capacities.

Time Interval [s]	Capacity [MW]				
	0.1	1	3	6	9.5
From 2 to 10	12%	6%	4%	3%	2%
From 10 to 30	12%	12%	9%	8%	6%
From 30 to 120	8%	20%	22%	22%	22%

4.4 Final considerations

In this chapter, a comparison among different layouts of PV power plant has been performed. It has been found that, with equal total capacity C, a distributed solution

provides a more stable output power than power plant composed by a single PV system located in a single location. This kind of results agrees with the one found on a real power plant with a capacity of 1.2 MW located in California which measurements are shown in [2] and reported in Fig. 43. It can be seen that, even though the distributed power plants are not perfectly fractionated in equal parts (like assumed in this thesis) as the number of subsystems increases, the variability decreases in according with all metrics, the same behavior that it has been observed in the simulated power plants with the same capacity.

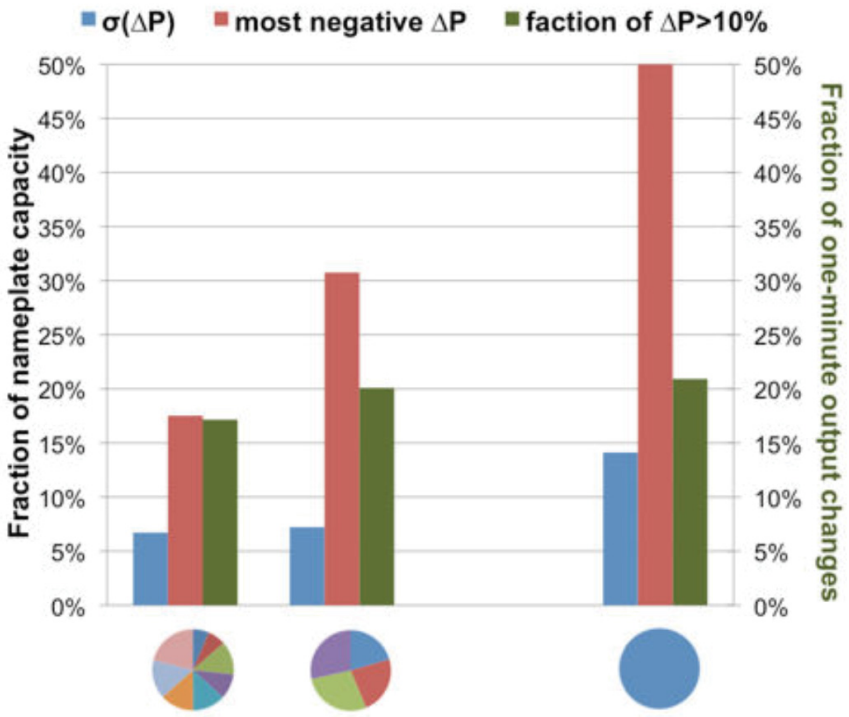

Fig. 43: Power output variability metrics for three power plant with the same total capacity of 1.2 MW [2].

A good agreement has been found also with the results shown in Table 5 at the end of the paragraph 2.3.3 where a reduction of variability is highlighted as the number of subsystems increases with the same total capacity.

Regarding the Time Interval, it has been confirmed that as the Time Interval increases, the standard deviation increases too, this result is in line with the one shown in Table 5, even if different timescales have been considered.

5 Conclusions

The present Master Thesis, which have concerned the smoothing effect of the output power variability on many kinds of PV power plant, proposes a method to simulate the output power of a distributed power plant, by knowing the irradiance point measurement data, the moving clouds speed and the system size. The experimental data measured in Tampere have been used and a distributed power plant composed by uncorrelated subsystems of the same capacity has been supposed in all performed simulations.

In the developed method, the smoothing effect due to the physical size and to the geographical dispersion has been modelled and therefore, the output power variability can be precisely estimated.

The proposed method, successively, has been validated, by comparing it with another one (presented in [25]) which, substantially, models the same effect and a good agreement has been found.

Moreover, an equivalence has been found in the calculation of the reduction of variability from a single subsystem to the entire PV fleet, when the power data are used instead of the irradiance data, this highlights that the smoothing effect due to the physical size of each system can be neglected in the evaluation of such reduction.

At the end, a comparison among different scenarios for a PV power plant has been performed in order to understand which is the solution that provides the more stable output power, taking in consideration the relation with the capacity of the power plant and the number of subsystems which composes the power plant. Generally, the results point out that, focussing on the difference between a centralized and a distributed solution, a distributed approach, with equal capacity, is more advantageous.

6 References

[1] J. Redmund, C. Calhau, L. Perret and D. Marcel, “Characterization of the spatio-temporal variations and ramp rates of solar radiation and PV,” 2015.

[2] A. Golnas and S. Voss, “Power Output Variability of PV System Fleets in three utility service territories in New Jersey and California,” 2010.

[3] J. Widén, “Variability and smoothing effects of PV power production: A literature survey”.

[4] A. Golnas, J. Bryan and G. Aghatehrani, “11-Month Power Output Variability Study of a PV system fleet in New Jersey,” in *26th European Photovoltaic Solar Energy Conference and Exhibition*, 12500 Baltimore Ave, Beltsville MD 20705, USA, 2011.

[5] A. Golnas and J. Bryan, “Output Variability in Europe's largest PV power plant,” 2011.

[6] A. Kankiewicz, D. Moon and M. Sengupta, “Observed impacts of transient clouds on Utility-scale PV fields,” 2010.

[7] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira and E. Izco, “Power output fluctuations in large scale PV plants: one year observations with one second resolution and a derived analytic model,” 2010.

[8] J. Marcos, L. Marroyo, E. Lorenzo and M. Garcia, “Smoothing of PV power fluctuations by geographical dispersion,” 2011.

[9] A. Murata, H. Yamaguchi and K. Otani, “A Method of Estimating the Output Fluctuation of Many Photovoltaic Power Generation Systems Dispersed in a Wide Area,” 2009.

[10] E. Wiemken, H. G. Beyer, W. Heydenreich and K. Kiefer, “Power Characteristics of PV Ensembles: Experiences from the Combined Power Production of 100 Grid Connected PV Systems Distributed Over the Area fo Germany,” 2001.

[11] K. Otani, J. Minowa and K. Kurokawa, “Study on areal solar irradiance for analyzing areally-totalized PV systems,” 1997.

[12] N. Kawasaki, T. Oozeki, K. Otani and K. Kurokawa, “An evaluation method of the

fluctuation characteristics of photovoltaic systems by using frequency analysis," 2006.

[13] M. Lave and J. Kleissl, "Solar variability of four sites across the state of Colorado," *Rewenable Energy,* vol. 35, pp. 2867-2873, 2010.

[14] A. Mills and R. Wiser, "Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power," 2010.

[15] M. Lave, J. Kleissl and E. Arias-Castro, "HIgh-frequency irradiance fluctuations and geographic smoothing," *Solar Energy,* vol. 86, pp. 2190-2199, 2011.

[16] A. Woyte, R. Belmans and J. Nijs, "Fluctuations in instantaneous clearness index: Analysis and statistics," *Solar Energy,* vol. 81, pp. 195-206, 2007.

[17] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira and E. Izco, "From irradiance to output power fluctuations: the pv plant as a low pass filter," *Progress in Photovoltaics: Research and applications,* no. 19, pp. 505-510, 2011.

[18] M. Lave, J. Kleissl and J. Stein, "A Wavelet-Based Variability Model (WVM) for Solar PV Power Plants," *IEEE Transactions on Sustainable Energy,* vol. 4, no. 2, pp. 501-509, 2012.

[19] P. Ineichen, "A broadband simplified version of the Solis clear sky model," 2008.

[20] M. Lave and J. Kleissl, "Cloud speed impact on solar variability scaling – Application to the wavelet variability model," *Solar Energy,* vol. 91, pp. 11-21, 2013.

[21] T. Hoff and R. Perez, "Quantifying PV Power Output Variability," *Solar Energy,* vol. 84, pp. 1782-1793, 2009.

[22] K. Lappalainen and S. Valkealahti, "Recognition and modelling of irradiance transitions caused by moving clouds.," *Solar Energy,* vol. 112, pp. 55-67.

[23] K. Lappalainen and S. Valkealahti, "Analysis of shading periods caused by moving clouds," *Solar Energy,* vol. 135, pp. 188-196.

[24] T. Hoff and R. Perez, "Modeling PV fleet output variability," *Solar Energy,* vol. 86, pp. 2177-2189, 2012.

[25] J. Marcos, I. de la Parra, G. M. and M. L., "Simulating the variability of dispersed

large PV plants," *PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS,* no. 24, p. 680–691, 2016.